

JOINT EVALUATION OF GENERAL BUDGET SUPPORT 1994–2004

Burkina Faso, Malawi, Mozambique, Nicaragua, Rwanda, Uganda, Vietnam



Final Inception Report

20 May 2005



Development
Researchers'
Network



International Development Department
School of Public Policy
University of Birmingham
Birmingham B15 2TT, U.K.
Tel: +44 (0) 121 414 3120
Fax: +44 (0) 121 414 4989
Website: www.idd.bham.ac.uk

Study contacts

Team Leader: Stephen Lister slister@mokoro.co.uk
Coordinator: Rebecca Carter r.l.carter@bham.ac.uk

Joint Evaluation of General Budget Support

FINAL INCEPTION REPORT

20 May 2005

Contents

Acronyms	iii
1. Introduction	1
Purpose and Scope of the Evaluation	1
Purpose of the Inception Report	2
Inception Report Structure	3
2. Perspectives on Partnership GBS	5
Introduction	5
The Definition of General Budget Support	5
The Origins and Characteristics of Partnership GBS	5
Broad Types of GBS	5
Origins and Purposes of Partnership GBS	7
Implications for Evaluation	8
Overview of GBS in the Case-Study Countries	8
Approach	8
Relevant Characteristics of the Case-Study Countries	9
The GBS Inventory	14
Implications for the Evaluation of GBS	15
3. Approach and Methods – Enhanced Evaluation Framework	19
Introduction	19
Challenges in GBS Evaluation	19
Complexity	19
Basic Evaluation Principles	20
Causality and Attribution	21
Strengths and Limitations of the Evaluation Framework	22
Origins and Outline of the Evaluation Framework	22
TOR Requirements for Refinement of the EF	23
The IDD Approach	25
The Enhanced Evaluation Framework	26
Assessment of the Evaluation Framework	26
Overview of the Enhanced Evaluation Framework	27
Using the EEF	28
Additional Methodological Issues	30
Standard Terminology	30
Counterfactuals	30
4. Key Evaluation Questions	33
Introduction	33
Main Causality Hypotheses and Detailed Evaluation Questions	33
Focusing and Prioritising the Evaluation	36
Principal Causality Chains and Causality Map	36
Key Evaluation Questions	37
Structure for Country Reports	37
Cross-Cutting Issues and Themes	40
Limitations of the Evaluation	41

5. Work Plan	43
Introduction	43
Overall Timetable	43
Quality Assurance and Comparability	45
Internal Quality Assurance	45
Ensuring that the Country Studies can be Synthesised	45
Stakeholder Consultation	46
Field Studies and the Country Reports	46
The Synthesis Note	48
The Synthesis Report	49
The Note on Approach and Methods	50
Boxes	
Box 1.1: Inception Period: Key Events and Dates	2
Box 1.2: Guide to the Annexes	4
Box 2.1: General Definition of Budget Support and GBS	6
Box 2.2: A Stylised Classification of GBS	6
Box 2.3: Feasibility of an Econometric Approach	8
Box 3.1: The DAC Evaluation Criteria	20
Box 4.1: Enhanced Evaluation Framework – Logical Sequence of Effects	34
Box 4.2: Key Evaluation Questions	39
Box 4.3: Chapter Structure of Country Reports	40
Box 5.1: Main Study Phases and Activities, Post-Inception	43
Box 5.2: Main Steps in Field Studies and Country Report Preparation	47
Box 5.3: The Synthesis Note Stage	48
Box 5.4: The Synthesis Report Stage	49
Figures	
Figure 3.1: GBS Evaluation Framework (simplified version)	24
Figure 3.2: The Enhanced Evaluation Framework (schematic view)	29
Figure 4.1: Causality Map for the Enhanced Evaluation Framework	38
Figure 5.1: Revised Timetable	44
Annexes	
Annex A: Terms of Reference	55
Annex B: Structure and Organisation of the Study	79
Annex C: Original Evaluation Framework (Executive Summary)	87
Annex D: Bibliography	95
Annex E: GBS Inventory in Study Countries	119
Annex F: Terminology	199
Annex G: Enhanced Evaluation Framework (Detailed Questions)	213
Annex H: Cross-Cutting Issues	229
Annex I: Country Field Studies	251
Annex J: Country Report Structure	257
Annex K: Key Evaluation Questions	265

Acronyms

AAP	Assessment and Action Plan
APR	Annual Progress Report
ARV	Anti-Retroviral
BoP	Balance of Payments
BS	Budget Support
CABS	Common Approach to Budget Support
CB	Capacity Building
CCI	Cross Cutting Issues
CDC	Community Development Committee
CDF	Comprehensive Development Framework
CEPEX	Central Projects and External Finance Bureau
CFAA	Country Financial Accountability Assessment
CG	Consultative Group
CIDA	Canadian International Development Agency
CoB	Central Bank (Mozambique)
COM	Council of Ministers
CPAR	Country Procurement Assessment Report
CPRGS	Comprehensive Poverty Reduction and Growth Strategy
CPV	Communist Party of Vietnam
CR	Country Report
CRS	Credit Reporting System
CSLP	Cadre Stratégique de Lutte contre la Pauvreté (Burkina Faso)
CSO	Civil Society Organisation
CSR	Civil Service Reform
CT	Country Team
CTL	Country Team Leader
DAC	Development Assistance Committee (of OECD)
DCI	Development Cooperation Ireland
DFID	Department for International Development (UK)
DIC	Department for International Funds (Mozambique)
DNPO	National Directorate of Plan and Budget (Mozambique)
DP	Development Partner
DRC	Democratic Republic of Congo
EC	European Commission
ECM	Economic Council of Ministers (Mozambique)
EEF	Enhanced Evaluation Framework
EF	Evaluation Framework
EIA	Environmental Impact Analysis
EIU	Economist Intelligence Unit
EQ	Evaluation Question
ESAF	Enhanced Structural Adjustment Facility
ESP	Economic and Social Plan
FCFA	Franc of the African Financial Community
FRDP	Fiscal Restructuring and Deregulation Programmes (Malawi)
G16	Group of Development Partners providing GBS to Mozambique
GBS	General Budget Support
GDP	Gross Domestic Product
GFATM	Global Fund for AIDS, Tuberculosis and Malaria
GNI	Gross National Income

GoM	Government of Mozambique
H&A	Harmonisation and Alignment
HIPC	Heavily Indebted Poor Countries
IADB	Inter-American Development Bank
IDD	International Development Department (University of Birmingham)
IFI	International Financial Institution
IMF	International Monetary Fund
IMR	Infant Mortality Rate
INSD	National Bureau of Statistics and Demography (Burkina Faso)
IP	International Partner
I-PRSP	Interim Poverty Reduction Strategy Paper
IR	Inception Report
IR1	First Draft Inception Report ¹
IR2	Revised Inception Report ²
JA	Joint Agreement
JP	Joint Programme
LENPA	Learning Network on Program-Based Approaches
LG	Local Government
LTEF	Long Term Economic Framework
M&E	Monitoring and Evaluation
MADER	Ministry of Agriculture (Mozambique)
MDG	Millennium Development Goal
MFPED	Ministry of Finance, Planning and Economic Development
MG	Management Group
MIDED	Ministry of Education (Mozambique)
MISAU	Ministry of Health (Mozambique)
MOU	Memorandum of Understanding
MPF	Ministry of Planning and Finance
MTBF	Medium Term Budget Framework
MTEF	Medium Term Expenditure Framework
n.a.	not available
NDP	National Development Plan
NEMA	National Environmental Management Agency
NEPAD	New Partnership for Africa's Development
NGO	Non-Governmental Organisation
Norad	Norwegian Agency for Development Cooperation
NRM	National Resistance Movement (Uganda)
ODA	Official Development Assistance
ODI	Overseas Development Institute
OE	State Budget (Mozambique)
OECD	Organisation for Economic Cooperation and Development
OHCHR	Office of the High Commissioner for Human Rights
PAF	Poverty Action Fund (Uganda) / Performance Assessment Framework (elsewhere)
PAP	Project Aid Partners
PARPA	Plano de Acção para a Redução da Pobreza Absoluta (Mozambique's PRSP)
PBA	Programme Based Approach
PE	Public Expenditure
PEAP	Poverty Eradication Action Plan
PEFA	Public Expenditure and Financial Accountability

¹ *Joint Evaluation of Budget Support – Inception Report, First Draft*, January 2005. Birmingham: IDD.

² *Joint Evaluation of Budget Support – Revised Inception Report*, March 2005. Birmingham: IDD.

PES	Plano Economico e Social (Economic and Social Plan) (Mozambique)
PFM	Public Finance Management
PGBS	Partnership General Budget Support
PIP	Public Investment Programme
PIU	Project Implementation Unit
PM	Prime Minister
PND	Plan Nacional de Desarrollo (National Development Plan)
PPA	Participatory Poverty Assessment
PR	Poverty Reduction
PRGF	Poverty Reduction and Growth Facility
PROAGRI	Sector Wide Approach to the Ministry of Agriculture (Mozambique)
PRS	Poverty Reduction Strategy
PRSC	Poverty Reduction Support Credit
PRSP	Poverty Reduction Strategy Paper
PSR	Public Sector Reform
PTA	Parent Teacher Association
PWT	Penn World Tables
QA	Quality Assurance
ROPE	Results Oriented Public Expenditure
RSA	Republic of South Africa
SAF	Structural Adjustment Facility
SAI	Supreme Audit Institution
SB	State Budget
SBS	Sector Budget Support
SECO	State Secretariat for Economic Affairs (Switzerland)
SG	Steering Group
Sida	Swedish International Development Cooperation Agency
SISTAFE	State Integrated Financial Management System (Mozambique)
SME	Small and Medium Scale Enterprise
SN	Synthesis Note
SOE	State Owned Enterprise
SPA	Strategic Partnership with Africa
SR	Synthesis Report
ST	Synthesis Team
SWAp	Sector Wide Approach
TA	Technical Assistance
TFDP	Task Force on Donor Practices
TL	Team Leader
TOR	Terms of Reference
U5MR	Under Five Mortality Rate
USAID	United States Agency for International Development
USD	United States Dollar
WAEMU	West African Economic and Monetary Union
WB	World Bank
WDI	World Development Indicators
WTO	World Trade Organisation

1. Introduction

Purpose and Scope of the Evaluation

1.1 This study, commissioned jointly by donors in collaboration with partner governments, is the first systematic cross-country evaluation of General Budget Support (GBS). The full Terms of Reference (TOR) are in Annex A. Details of the management structure for the study and the study team are in Annex B.

1.2 As summarised in the TOR (§3.1):

The purpose of the evaluation is to evaluate to what extent, and under what circumstances (in what country contexts), GBS is relevant, efficient and effective for achieving sustainable impacts on poverty reduction and growth. The evaluation should be forward looking and focused on providing lessons learned while also addressing joint donor accountability at the country level.

1.3 The evaluation is to cover the implementation and the results of GBS during the period 1994–2004 (TOR §3.3). The seven case-study countries are Burkina Faso, Malawi, Mozambique, Nicaragua, Rwanda, Uganda and Vietnam. These have been selected as an illustrative rather than representative sample of countries that receive GBS. Moreover, the particular focus of the evaluation is on "partnership GBS" ("new GBS" and "poverty reduction GBS" are equivalent terms). The characteristics and evolution of partnership GBS, generally and in the study countries, are discussed in Chapter 2. In all cases, however, partnership GBS was an innovation in the latter part of the review period. The study is not required to evaluate the forms of GBS that preceded partnership GBS, but it is expected to use the earlier experiences of programme aid as a point of comparison in assessing partnership GBS.

1.4 Eventual outputs of the study will include seven free-standing country-level Evaluation Reports, a Synthesis Note (SN) summarising the findings from the seven country studies, a final Synthesis Report (SR) which will extend the scope of the evaluation's conclusions, and a Note on Approach and Methods.³

1.5 The study is based on, but is also required to improve, a specially developed Evaluation Framework (EF). Annex C is the Executive Summary of the EF (Booth and Lawson, 2004).⁴

1.6 This report draws on a series of inception visits to the study countries (details in Annex B, summary timetable in Box 1.1 below). A first draft inception report (IR1),⁵ was discussed at a Steering Group (SG) meeting on 1–2 February 2004. This was thoroughly re-drafted in the light of comments and a revised inception report (IR2)⁶ was submitted in March. The present Final Inception Report responds to comments on the previous drafts. The main substantive changes are in Chapter 4.

³ The revised timetable for delivering these outputs is described in Chapter 5.

⁴ Booth, D. and Lawson, A. (2004). *Evaluation Framework for General Budget Support*. London: ODI. Throughout this report, Evaluation Framework (with initial capitals) always refers to the EF developed by Booth and Lawson. For the complete list of all references, please see Annex D.

⁵ IDD, *Joint Evaluation of Budget Support – Inception Report, First Draft, January 2005*.

⁶ IDD, *Joint Evaluation of Budget Support – Revised Inception Report, March 2005*.

Box 1.1: Inception Period: Key Events and Dates

Key Events	2004
Contract Start Date (mobilisation)	August 23
Team Workshop	September 13–14
Contract Signature	October 15
Country Inception Visits	October–December
<i>Vietnam</i>	<i>Oct 10–23</i>
<i>Mozambique</i>	<i>Oct 16–30</i>
<i>Rwanda</i>	<i>Oct 31–Nov 16</i>
<i>Malawi</i>	<i>Nov 1–12</i>
<i>Nicaragua</i>	<i>Nov 1–14</i>
<i>Uganda</i>	<i>Nov 14–Dec 2</i>
<i>Burkina Faso</i>	<i>Nov 28–Dec 13</i>
	2005
Draft Inception Report	January 24
SG meeting to discuss Draft Inception Report	February 1–2
Team workshop	February 23–24
Revised Draft Inception Report	March 14
Team Workshop	April 26–27
Final Inception Report	May 20

Purpose of the Inception Report

1.7 The TOR (§6) specify the Inception Report as follows:

An Inception Report shall be submitted at the end of the inception phase providing details of the proposed approach and methods and specifying the issues and themes to be studied. The Inception Report shall in particular account for:

- *The findings of the first visits to the case-study countries, including an overview of GBS support and possible consequences for the focus, approach and methodology of the evaluation.*
- *The application of the [Evaluation Framework], including the developed causality tree.*
- *The precise focus of the evaluation, i.e. the key issues and themes to be evaluated specifying any limitations of the evaluation.*
- *The approach and method, including data gathering and analysis (causality analysis and triangulation), as well as a specification of indicators and identification of key stakeholder groups.*
- *The approach to assessing the cross-cutting issues.*
- *The approach to ensure that the different case-study countries can be easily synthesised.*
- *The approach to ensure quality assurance throughout the country evaluations and the synthesis evaluation.*

- *The detailed work plan, specifying the organisation and time schedule for the evaluation process – allowing sufficient time for consultation with the SG, MG and other key stakeholders.*

Inception Report Structure

1.8 This Inception Report responds to these requirements as follows:

Chapter 2 describes the origins and distinguishing features of the partnership GBS that is the principal focus of this evaluation; it then provides an overview of partnership GBS in the case-study countries. (A more extensive inventory of GBS in these countries appears in Annex E.)

Chapter 3 reviews the GBS Evaluation Framework and other key aspects of the approach and methods for the study, and explains the Enhanced Evaluation Framework (EEF) that has been developed during the inception phase.

Chapter 4 draws on the EEF and on the findings from the inception visits to the case-study countries to propose the key evaluation questions to be addressed in the Country Reports (the study content). As anticipated, this is the chapter that has developed most since the submission of the Revised Inception Report.

Chapter 5 describes the work plan for the remainder of the study (the study process), including measures for quality assurance and to ensure comparability of findings across countries.

1.9 These five chapters comprise the concise main part of this Inception Report. They are supported by Annexes (see Box 1.2) which include additional background material and initial findings, along with other material that is principally intended as detailed guidance for the study teams undertaking the field phase. Annexes G, J and K (new) have been substantially revised since IR2.

Box 1.2: Guide to the Annexes

Title	Comment
A. Terms of Reference	Full TOR (25 pages).
B. Structure and Organisation of the Study	Details of Steering Group and Management Group; country contacts; the IDD consortium; team members and responsibilities; timetable to date; study website.
C. Evaluation Framework (Executive Summary)	Concise version of the Evaluation Framework for GBS – a methodological guide that forms an adjunct to the TOR.
D. Bibliography	Part 1 collects general references for the study. Part 2 highlights relevant cross-country studies in which the GBS case-study countries feature.
E. GBS Inventory in Study Countries	A working summary of data collected by country teams on GBS and other relevant aid programmes in each study country for 1994–2004.
F. Terminology	Standard definitions for key terms used in the evaluation.
G. Enhanced Evaluation Framework (Detailed Questions)	Detailed logical framework developed during the inception phase (to supersede the version associated with the original Evaluation Framework).
H. Cross-Cutting Issues	Explains the treatment of the four cross-cutting issues identified in the TOR (gender equality, environment, democracy & human rights, and HIV/AIDS).
I. Country Field Studies	Teams, timing and particular areas of focus for the country field studies.
J. Country Report Structure	Detailed guidelines for the format of the Country Reports.
K. Key Evaluation Questions	Key evaluation questions with guidelines on judgement criteria, causality chains, counterfactuals, relevant evidence and data sources.

2. Perspectives on Partnership GBS

Introduction

2.1 This evaluation focuses on General Budget Support (GBS) as a particular form of programme aid. A defining characteristic of GBS is that it provides unearmarked funds to the government budget, but it is not the only form of programme aid that does this, and it is distinguished from these other forms of programme aid by the purposes and the manner of the transfer. The current form of GBS took shape in the late 1990s, and is variously known as "new GBS", "poverty reduction GBS" and "partnership GBS". We have adopted the latter term as the one that best captures the intended change in aid relationships that this form of GBS embodies.

2.2 This chapter (a) explains the definition of GBS, (b) notes the origins and objectives of partnership GBS, and (c) provides an overview of GBS in the study countries.

The Definition of General Budget Support

2.3 Box 2.1 shows the definition of GBS, and its relationship to other forms of programme aid, as specified in the TOR (Annex A). A budget support programme comprises not only the funds themselves, but also the policy dialogue and the conditions linked to their disbursement, and related technical assistance and capacity building. The Evaluation Framework extends this further to include as additional inputs the alignment and harmonisation activities of the GBS donors.

2.4 Other forms of programme aid (including debt relief and other balance of payments support) may have similar effects in terms of generating resources that can be used to finance the government budget (and could therefore also be legitimately described as budget support), but different types of budget support are distinguished on the basis of the intent of the programme and on the accompanying inputs. An historical perspective is therefore essential to understand what distinguishes partnership GBS from the earlier forms of programme aid.

The Origins and Characteristics of Partnership GBS

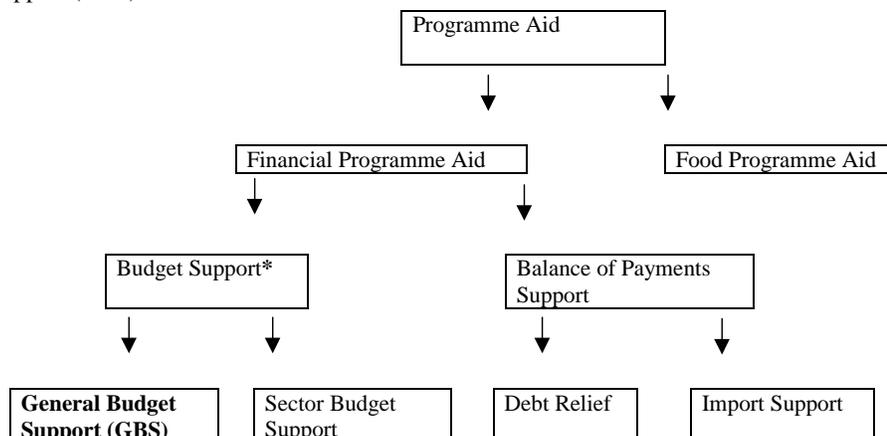
Broad Types of GBS

2.5 Programmatic (non-project) forms of aid have a very long history. The dominant form of financial programme aid during the 1980s and 1990s was structural adjustment finance provided mainly by International Financial Institutions (IFIs). In many ways (see below), partnership GBS has been a reaction to the perceived shortcomings of structural adjustment GBS, but both may be contrasted with unconditional GBS, as shown in Box 2.2.

2.6 The three broad types of GBS, depicted in Box 2.2, each consist of concessional finance to government that is not earmarked for a particular use, plus the conditions attached to this finance. It is the difference in attached conditions, linked to the purpose of the non-earmarked transfers, which distinguishes the three types. Distinguishing these broad types of GBS helps to resolve the paradox that some major donors of non-earmarked financial transfers to governments during the decade of our study period have also been sceptics regarding partnership GBS.

Box 2.1: General Definition of Budget Support and GBS

As defined for the purpose of this evaluation, **programme aid** can be divided into food aid and financial programme aid. Financial programme aid includes both budget support and balance of payments support (such as debt relief and import support). Budget support in turn can be divided into sector budget support and general budget support (GBS).



* Referred to as direct budget support in the *Evaluation Framework*.

The general characteristics of **budget support** are that it is channelled directly to partner governments using their own allocation, procurement and accounting systems, and that it is not linked to specific project activities. All types of budget support include a lump sum transfer of foreign exchange; differences then arise on the extent of earmarking and on the levels and focus of the policy dialogue and conditionality.

Sector Budget Support is distinguished from **General Budget Support** by being earmarked to a discrete sector or sectors, with any conditionality relating to these sectors. Additional sector reporting may augment normal government accounting, although the means of disbursement is also based upon government procedures.

Source: TOR and Evaluation Framework. For more detailed explanation of the terms involved (e.g. the distinction between real and notional earmarking) see Annex F (Terminology).

Box 2.2: A Stylised Classification of GBS

(non-earmarked concessional finance in all cases)

Type	Purpose	Conditionality
1. Unconditional GBS	To fill a temporary gap in public finances.	No specific conditions, often related to broader political, historical and trade ties, e.g. support to countries attaining independence (1960s-70s), US Marshall Aid to Europe (1940s), Japanese assistance to neighbours after Asian financial crisis (1990s). Essentially bilateral.
2. Structural Adjustment GBS	To reduce external and internal imbalances and raise economic growth.	Specific conditions focused on deregulation, privatisation, inflation and public sector deficits, e.g. numerous structural adjustment programmes. Led by IFI credits with bilateral support.
3. Partnership GBS <i>(also known as 'poverty reduction GBS' and 'new GBS')</i>	To raise the capacity of government to attain agreed poverty reduction targets.	Specific conditions related to performance in governance, service delivery to the poor, as well as inflation and public sector deficits. Led by IFI credits with bilateral support.

Origins and Purposes of Partnership GBS

2.7 The purposes, and the approach to conditionality, of partnership GBS are rooted in perspectives on the role and effectiveness of aid that emerged in the latter half of the 1990s. These derived from the following interacting impulses:

- Revised assessments of aid effectiveness, which argued that aid works in good policy environments but that necessary reforms cannot be bought through conditionality (Dollar et al, 1998, *Assessing Aid*,⁷ was seminal). Hence an emphasis on ownership, and on being more discriminating in allocating aid to countries and governments that are able to use it effectively.
- Growing concerns for debt relief and a more direct focus on poverty. Hence the direct link between the Heavily Indebted Poor Countries (HIPC) initiative and Poverty Reduction Strategy Papers (PRSPs), and the focus on the adequacy of government systems to allocate HIPC resources to pro-poor expenditures. Hence also the adoption of the Millennium Development Goals (MDGs). Although debt relief is not formally part of what is being evaluated (see Box 2.1), it was extremely important in developing the PRSP framework for GBS, while, for many bilateral donors, debt relief funding was a direct precursor of GBS.⁸
- Greater scepticism about standard policy prescriptions and more recognition of differences in country environments. Allied to the increased concern for ownership, this creates a greater concern for aid to support effective *processes* – not just the particular policies espoused and budgets announced, but the underlying systems for macroeconomic management, planning and budgeting, and associated systems of internal accountability. Again this is epitomised in the PRSP approach.
- A perception (also highlighted in the *Assessing Aid* literature) that inappropriate aid modalities had become part of the problem. Efforts to bypass weaknesses in government systems were seen to have further weakened them, to have fragmented national decision making, and to have raised the transaction costs of aid. Hence efforts to support greater coordination and harmonisation of aid reflected, inter alia, in the Comprehensive Development Framework (CDF), work by the OECD's Development Assistance Committee (OECD DAC), the Strategic Partnership with Africa (SPA) and other bodies on harmonisation and alignment, the trend towards Sector Wide Approaches (SWAs) and the adoption of the principles of Programme Based Approaches (PBAs).⁹ PBA principles (which apply equally to PRSPs, to sector approaches and to lower levels of engagement) ownership, alignment and harmonisation, the use of government systems, and so forth.¹⁰ Concern to disburse aid through government systems increases the focus on the quality of public finance management as well as on sound macroeconomic management.
- An additional factor is the Monterrey commitment to scale up aid substantially. In the light of the other concerns cited (harmonisation, ownership, the preferences of recipient governments), and because achieving MDGs may require the financing of recurrent costs, GBS is seen by its advocates as a particularly relevant modality for the channelling of increased aid flows.

⁷ Dollar, D. et al. (1998). *Assessing Aid: What Works, What Doesn't and Why*. New York: Oxford University Press for World Bank.

⁸ In terms of the stylised classification of Box 2.2, HIPC resources represent a form of unconditional budget support after the completion point is reached, since there is then an irrevocable flow of debt relief.

⁹ See Annex F (Terminology) for definitions.

¹⁰ As signatories to the Paris Declaration (*Paris Declaration on Aid Effectiveness*, High Level Forum, March 2005) donors have agreed that increased use of PBAs should be adopted as an indicator of progress in harmonisation.

2.8 Altogether, the range of expectations from GBS is extraordinarily wide. Its ultimate objectives may be relatively clear (at its simplest, poverty reduction – though there is not much that is simple about that!), but how this is expected to occur involves a much wider range of expectations. The TOR (§2.2) draw attention to:

- Improved *coordination* and *harmonisation* among donors and *alignment* with partner country systems (including budget systems and result systems) and policies.
- *Lower transaction costs*.
- *Higher allocative efficiency of public expenditures*.
- *Greater predictability* of funding (to avoid earlier “stop and go” problems of programme aid).
- *Increased effectiveness of the state and public administration* as GBS is aligned with and uses government allocation and financial management systems.
- *Improved domestic accountability* through increased focus on the Government’s own accountability channels.

Implications for Evaluation

2.9 Project evaluation methodology is well established, and relatively straightforward (although such evaluations frequently abstract from the problems of fungibility – see the discussion of counterfactuals in Chapter 3). PBAs are inherently much more complex because of their breadth, the provision of joint inputs, and the nature of their objectives. GBS, as the form of PBA with the broadest scope, is also the most challenging to evaluate. We take up these challenges in Chapter 3. First, we consider how trends towards partnership GBS have been manifested in the case-study countries.

Overview of GBS in the Case-Study Countries

Approach

2.10 Because of its scale, GBS is inherently lumpy, and there has been a finite set of budget support operations in each of the study countries between 1994–2004. Our approach has been to identify them all individually, and to build up a comprehensive inventory for each study country based on programme-level information. (The possibility of complementing country-level information with an econometric analysis of standard international data has been considered but rejected – see Box 2.3.)

Box 2.3: Feasibility of an Econometric Approach

The subtleties of the distinctions involved, the lack of standard definitions, and the novelty of partnership GBS as a major form of programme aid, mean that standard international data on aid are of limited value for charting trends in the types and levels of GBS flows.¹¹ As part of the inception phase of this study, the feasibility of an econometric analysis of GBS performance was explored.¹² Although there appeared to have been significant recent improvements in the principal (OECD DAC) data sources, it has been decided not to pursue this line of enquiry within the present study because of remaining doubts about the reliability of the data, because the proposed approach would not have yielded robust country-level conclusions (it would have incorporated the seven case-study countries within a wider 42-country panel), and because it could not have provided any insights on the differential performance of different designs within the category of partnership GBS.

¹¹ See Lanser, P. (2003). *Inventory of Programme Aid 1992 – 2001. Preparatory study for the planned joint evaluation of General Budget Support*. Rotterdam/The Hague: ECORYS-NEI (Working Document).

2.11 We have deliberately sought to capture all forms of budget support, not just what is unambiguously partnership GBS according to the TOR definition. This approach recognises that different donors have different terminology and definitions (partly for presentational reasons) and that identifying partnership GBS on the basis of donors' own classifications and programme labels would therefore be treacherous. Furthermore, even with detailed programme-level knowledge it is difficult to draw sharp distinctions between different types of programme aid and budget support. For example the difference between GBS and Sector Budget Support (SBS) – see Box 2.1 above and the fuller discussion in Annex F – can turn on a subtle interpretation of the difference between real and virtual earmarking. In practice there is a spectrum of related aid instruments, and the drawing of sharp boundaries between different types is likely always to be somewhat arbitrary.

2.12 In addition, many of the design elements of GBS also feature both in earlier forms of programme aid and in current co-existing modalities. Even though the focus of the present study is on partnership GBS, there is potential for useful insights from comparisons with other forms of programme aid, particularly when considering different design elements (e.g. conditionality, performance indicators) that are common to GBS and the earlier and other contemporary versions of budget support. The TOR endorse the pragmatic approach to field work that we have adopted, and it is reflected in the scope of information being collected for the inventory (Annex E).

2.13 Country contexts are different. Even countries which are quite similar in many dimensions may have a different pattern and history of the aid relationships out of which GBS has evolved. This will be an important theme to pursue in the country reports themselves, and in the synthesis phase of this study. For the purposes of this Inception Report, the next paragraphs briefly characterise some of the most striking features of the case-study countries, before we provide an overview of the inventory and its contents.

Relevant Characteristics of the Case-Study Countries

2.14 The most obvious grouping of the case-study countries is that the five African countries represent a relatively homogeneous group in terms of economic and social structure, state of development and patterns of development aid, while Nicaragua and Vietnam are clear outliers. But beyond this starting point there are distinct political and economic contexts by country:

- While all are classified as ‘poor’ countries, each country has distinct poverty levels and characteristics, and significantly different recent trends in poverty reduction.
- Each country also has a particular history of aid relationships, and of budget support intent and practice during the period 1994–2004, involving very different levels of experience in partnership GBS (over time, by donor plurality, and by the quantity and relative significance of GBS funds involved).

¹² Barassi, M. and Ercolani, M. Feasibility Study for Quantitative Analysis, Annex H of IDD (2005). *Joint Evaluation of Budget Support – Inception Report, First Draft, January 2005*. Birmingham: IDD. (IR1).

Burkina Faso

2.15 In 2003, 46.4% of the population were living below the poverty line,¹³ while in the same year Gross National Income (GNI) per capita¹⁴ was \$200. Burkina Faso is heavily dependent on Official Development Assistance (ODA): net ODA/GNI in 2003 was 10.8%.¹⁵ A democratic government has been in place since 1991. The last two presidential elections have been won by the present incumbent, who faces another election in 2005. Legislative elections in 2002 resulted in an enlarged opposition presence in parliament. Government-donor relations have at times been affected by governance issues (e.g. concerning arms-dealing activities and episodes of human rights abuse).

2.16 Two periods of aid relationship are identifiable during the study period 1994–2004:

- (i) the 1990s: characterised by structural adjustment programmes; and
- (ii) from 2000: elaboration and implementation of the CSLP (Burkina Faso's PRSP), with Burkina Faso applying for HIPC initiative debt relief.

2.17 GBS is provided through the World Bank's Poverty Reduction Support Credit (PRSC, now in its fourth phase since 2001) and a group of bilateral donors (Belgium, Denmark, EC, France, Netherlands, Sweden, Switzerland and others), who have been supporting the Poverty Reduction Strategy (PRS) since 2001 and have adopted the first protocol for joint budget support. At the same time, non-GBS donors (namely Canada, Germany, Japan and USAID) continue to be important aid providers.

Malawi

2.18 Over 60% of the population live below the poverty line.¹⁶ GNI per capita was \$170 in 2003. Malawi is heavily aid dependent: net ODA/GNI in 2003 was 29.5%. Malawi has been an elective democracy since 1994, but government-donor relations have been affected by the lack of fiscal discipline evident throughout 1994–2004. In this context, the new presidency as a result of the 2004 elections is seen as a potentially positive change.

2.19 Two periods of aid relationship are identifiable during the study period 1994–2004:

- (i) the 1990s: dialogue led by the IMF and the World Bank, focusing on trade liberalisation and structural reforms; and
- (ii) from 2000 onwards: transition to a focus on the IMF Poverty Reduction and Growth Facility (PRGF) agreement, with key bilateral donors (Denmark, EC, Norway, Sweden and UK) promoting a coordinated approach with the PRGF under the CABS (Common Approach to Budget Support) framework.

2.20 However, with a lack of fiscal discipline evident throughout the period, the PRGF was off track at the point of signing in 2000, and actual releases of GBS have been limited. There has been some mutual recrimination between government and donors as a result.

¹³ Poverty line defined as 82,672 FCFA in 2003. INSD household survey 2003.

¹⁴ All GNI per capita data are USD, Atlas method.

¹⁵ Source for GNI and ODA/GNI statistics for all seven case-study countries is OECD DAC, *Aid at a Glance* (http://www.oecd.org/countrylist/0,2578,en_2649_34447_25602317_1_1_1_1,00.html).

¹⁶ As defined by the WB.

Mozambique

2.21 53.6% of the population live below the poverty line.¹⁷ GNI per capita was \$210 in 2003. Mozambique is heavily aid dependent: net ODA/GNI in 2003 was 25.2%. Since the pre-1992 situation of civil turmoil and the severe disruption of government operations, there has been a democratic multi-party regime with the last elections held in December 2004. Governance issues have affected government–donor relations (particular crises have stemmed from corruption-related banking scandals and episodes of human rights abuse).

2.22 Three periods of aid relationship are identifiable during the study period 1994–2004:

- (i) 1992–1996: preparation for peace, with emergency and food aid dropping off sharply between 1993–94;
- (ii) 1996–2000: gradual decrease in multilateral and non-governmental organisation (NGO) roles, with growing donor participation in recurrent expenditure; and
- (iii) 2001–2004: introduction of partnership GBS.

2.23 Of the large numbers of donors operating in Mozambique, 17 of them now collaborate in offering budget support, with coordination of GBS formalised in 2000 in a joint donor programme. The PRSP (PARPA – Plano de Cacao Para a Redu e ad Pores Absolute – in Mozambique) approved in 2001 by IMF and the WB, has been followed with a common Performance Assessment Framework in 2003. Except for the USA, all the large donors and many of the smaller bilaterals now participate in the principal grouping, the Group of 16 (G16), that have joined together to provide general budget support. Even 'outsiders' like the USA and Japan participate as 'observers' in G16 meetings, and the volume and proportion of aid that is channelled through the budgetary process has increased consistently over the last seven years.

Nicaragua

2.24 In 2001, 45.8% of the population were living below the poverty line,¹⁸ although GNI per capita was \$730 in 2003 – much the highest among the case-study countries. Nicaragua is nevertheless also heavily aid dependent: net ODA/GNI in 2003 was 21.3%. Political life was extremely polarised between the Sandinista Front and the Liberal Alliance until the relatively recent pact between the parties' leaders. In the last decade the context of government–donor relations has been the economic fallout from natural disasters (Hurricane Mitch) and political crises (the prosecution of former government officials on charges of corruption).

2.25 Two periods of aid relationship are identifiable during the study period 1994–2004:

- (i) 1990s: the first IMF Enhanced Structural Adjustment Facility (ESAF) expired in 1997 without a single one of the three annual programmes keeping on track, while the second ESAF, in 1998, went off track the same year; and
- (ii) from 2000: dominated by the HIPC initiative (decision point 2001 and completion point 2004) and its requirement for a PRSP to be drawn up. The World Bank PRSC started in 2004 and there is now a vigorous effort by donors (led by bilaterals – Finland, Norway, Sweden, Switzerland and others) to agree a Joint Financing Arrangement on GBS with performance conditions based on the PND (Plan Nacional de Desarrollo).

¹⁷ As defined by INE/Republic of Mozambique 2004.

¹⁸ Standard of Living Survey 2001.

Rwanda

2.26 60.3% were living in poverty in 1999/2000.¹⁹ GNI per capita was \$220 in 2003. Rwanda is heavily aid dependent: net ODA/GNI in 2003 was 20%. The watershed marking Rwanda's modern history and its relationship with donors is the genocide of 1994. Immediately afterwards there was a period of extreme chaos in the 'aid market' with large numbers of donors active – multinationals, bilaterals and NGOs. Rwanda is currently governed by a democratic multi-party regime; elections were last held in 2003. Government–donor relations have been affected by the crisis prompted by the war with DRC in 1998 and then more recently by over-spending including on allegedly unjustified items (2003). A new crisis may be in the making with clashes in Eastern DRC (end 2004).

2.27 Three periods of aid relationship are identifiable during the study period 1994–2004:

- (i) 1994–1997/98: characterised by virtually no government accountability or consultation with the population;
- (ii) 1997/98–2002: the government (re-)shaping itself, the emergence of Vision 2020 closely followed by the preparation of the PRSP; and
- (iii) from 2002, when the PRSP was endorsed. There has been a progressive shift of budget support operations towards the partnership GBS paradigm, with the Rwanda Government strongly stating its wish to see more donors providing a greater proportion of aid through GBS. Key GBS donors are DFID, EC, Sida and the World Bank, among others. GBS and GBS partnerships are at a relatively early stage, and 2005 will be a test period, especially as the Government plans to undertake a thorough revision and updating of the PRSP in this year.

Uganda

2.28 Headcount income poverty²⁰ reduced from 56% in 1992 to 44% in 1997 and further to 35% in 2002; however this indicator increased to 38% in 2003. GNI per capita was \$240 in 2003. Uganda's present constitutional framework was established by the National Resistance Movement (NRM) government which came to power in 1986. Subsequent Presidential and Parliamentary elections on a non-party basis (1996 and 2001) have been won by the NRM, and Uganda has yet to experience a democratic change of government. The run-up to forthcoming presidential elections may well mark another phase in donor–government relations. This is likely to depend on a variety of factors, including progress towards multi-party democracy and whether the presidential term limits embodied in the present Constitution are adhered to, and the human rights record in the run-up to the elections.

2.29 Uganda is notable for the extent of government leadership in the poverty reduction strategy and in innovative approaches to aid management. Three periods of aid relationship are identifiable during the study period 1994–2004:

- (iv) 1994–1997 pre- PEAP (Poverty Eradication Action Plan)²¹: establishment of a stable macroeconomic environment and fiscal discipline through the introduction of cash budgeting and the Medium Term Budgetary Framework (MTBF) in 1994, following an earlier lapse in fiscal discipline. Liberalisation and privatisation policies quickly followed, sponsored by the World Bank and IMF, and supported by Structural Adjustment lending.

¹⁹ PRSC 2004.

²⁰ Based on the national poverty line.

²¹ The Poverty Eradication Action Plan was a Ugandan initiative and became the prototype of PRSPs.

- (v) 1997–2000 (PEAP I): implementation of the Poverty Eradication Action Plan, and the introduction of the Medium Term Expenditure Framework (MTEF) in 1997. During this time Uganda first benefited from HIPC debt relief, prompting the formation of the Poverty Action Fund (PAF) by government. The combination of MTEF, SWAps and the PAF facilitated donors' move towards the provision of budget support, much of which was notionally earmarked to the PEAP via the PAF. Towards the end of this period the second iteration of the PEAP was developed, with a deeper, more evidence-based and participatory process than the original PEAP. PEAP II acted as the Government of Uganda's PRSP and qualified Uganda for debt relief under the enhanced HIPC initiative.
- (vi) 2000–2004 (PEAP II): the period of implementation of the PEAP, and the development of the current modalities for General Budget Support, centred on the World Bank's PRSCs, the first of which was provided in 2001. Towards the end of this period the third iteration of the PEAP was finalised.

Vietnam

2.30 Vietnam experienced a remarkable growth-led decline in the incidence of poverty in the 1990s, from 58.2% in 1992/93 to 37.4% in 1997/98.²² GNI per capita was \$480 in 2003. However, if present trends continue, Vietnam will graduate to Middle Income Country status within the foreseeable future. Vietnam is not heavily aid dependent; net ODA/GNI in 2003 was 4.5%. The Doi Moi economic reforms began in 1986, but it was not until the lifting of the US embargo in 1994 that the International Financial Institutions (IFIs) were able to begin lending to Vietnam (policy advice not linked to funding had been given previously). The previously small group of donors in Vietnam is now much larger.

2.31 Three periods of aid relationship are identifiable during the study period 1994–2004:

- (i) 1994–97: the first Consultative Group meeting was held in 1994, followed by a decree on aid management in 1998;
- (ii) 1998–2000: as a response to the Asian crisis, further reforms were initiated and, from 1999, substantial budget support was provided by Japan through the Miyazawa initiative; in mid-2000 the preparation of an Interim Poverty Reduction Strategy Paper (I-PRSP) began, as a route to accessing PRGF and PRSC funding, although debt relief was not sought; and
- (iii) from 2001: the development of the Comprehensive Poverty Reduction and Growth Strategy.²³

2.32 The IMF PRGF programme began in 2001 but was suspended in 2002, while the WB PRSC I began in 2001/02, followed by PRSC II (2003) and PRSC III (2004). The original PRSC I was co-financed by the UK, Netherlands, Denmark and Sida. By PRSC IV this had expanded to a more diverse group of potential co-financiers including Japan, France and the EC. At the same time, the government continues to be comfortable with a project-focused approach. This mirrors its own approach to implementation, involving discrete capital projects run by project management units which are semi-integrated into government systems; partnership GBS remains a relatively small, as well as recent, component of aid flows.

²² Glewwe P, Gragnolati M, and Zaman H. (2002) *Who gained from Vietnam's boom in the 1990s? An analysis of Poverty and Inequality Trends*. Economic Development and Cultural Change (volume 50).

²³ Comprehensive Poverty Reduction and Growth Strategy (Vietnam's PRSP).

The GBS Inventory

2.33 Our inventory of GBS (summarised in Annex E)²⁴ attempts to capture the main features of successive general budget support (and closely related) programmes in each case country. The primary purpose of the inventory is to identify programme intentions, procedures and inputs so that their outputs and outcomes can then be examined in the evaluation. By taking a broad view of the programmes that are relevant to this evaluation, it also supports the historical and comparative perspective that has been stipulated (¶1.3 above). Moreover, as we discuss in Chapter 3, entry conditions for GBS and the interactions among donors and among aid modalities, both at input level and prior to GBS commencement, have emerged as important subjects for this study, and the inventory helps in characterising these.

2.34 Apart from the identification of the different international partners involved in each programme and the scale of assistance involved, the inventory seeks to capture the following aspects of the programmes:

- (a) **Programme intent:** as we have noted, the intention behind the provision of unearmarked budget support is one of the principal distinguishing features of partnership GBS. The inventory seeks to capture changing intentions over time as well as the differences in emphasis between donors simultaneously involved.
- (b) **Alignment with national strategies:** alignment with national goals and strategy for poverty reduction is fundamental to partnership GBS.
- (c) **Disbursement procedures:** disbursement via government systems is a defining characteristic of GBS, but other aspects of the transfer (tranching, conditions and predictability of disbursement) may have an important bearing on the consistency between the intentions behind GBS and the form that it takes.
- (d) **Conditionality and performance indicators:** as already noted, partnership GBS is ostensibly based on a different approach to conditionality than under structural adjustment programmes; it is important to review whether the evidence bears out claims that the nature of the government–donor relationship has changed, and to what extent. The indicators linked to a programme are directly related to its conditionality, and are also very revealing about the intent of the programme. We attempt to capture such aspects as the number of indicators, and their nature (e.g. whether they are process or results oriented, whether they are drawn from the partner country's PRS, whether they are linked to performance indicators for sector programmes and so forth).
- (e) **Links to technical assistance (TA) and capacity building:** whether the programme content is oriented to capacity building in core government services, and whether the programmes are explicitly linked to TA and capacity building inputs, e.g. for strengthening public finance management.
- (f) **Procedures for dialogue:** the structure and content of dialogue for partnership GBS is supposedly characterised by orientation to government leadership, capacity building and long-term commitment to poverty reduction. We seek to identify the specific arrangements for dialogue related to GBS programmes, set in the context of pre-existing and wider institutions for interaction between the government and its international partners.

²⁴ Annex E is a snapshot of work in progress. The information it summarises will be reviewed and revised during the preparation of Country Reports; refined inventory information will be incorporated in the Country Reports themselves.

- (g) **Donor harmonisation and alignment (H&A):** within partnership GBS, donor harmonisation and alignment with recipient country systems, as well as policies, are regarded as essential for increased ownership by government, and lower costs and greater effectiveness of core government services. As with dialogue arrangements, we seek to locate GBS-related H&A in a wider context and to understand their evolution.
- (h) **Experience in implementation:** perceptions, as well as documented reviews and evaluations, are important here, since it is clear that GBS has evolved, and continues to evolve, on the basis of learning and interpretation of past experiences.
- (i) **Sources of evidence:** these are important to record for further reference and additional detail; also, in some cases, as important secondary source material on the main themes of the present evaluation; and further, as a means of checking on possible biases in the sourcing of information.

2.35 This last point is an important one. Programme by programme, donor records are more systematic than governments' records tend to be. However, attempting to reconcile financial data, in particular, between donor and government sources is both difficult and revealing. Donors' records regularly indicate higher aid flows than governments register, and, despite the aspirations of budget support to align with government systems, it is extremely difficult to get reliable, donor-sourced disbursement information that matches governments' fiscal years and budgetary classifications. Equally, donors are more systematic than governments in reviewing and evaluating such programmes, and it is important to triangulate donor-sourced information with the experiences of other stakeholders. Written records are biased to the formal and intended dimensions of GBS, whereas the way GBS operates informally and in practice may show substantial variance. Nevertheless, the information assembled in this inventory is an essential base for the GBS evaluation. Country by country, it provides evidence for constructing the history and establishing the current dimensions and likely future trends in GBS to be set in the context of aid flows and relationships more generally.

Implications for the Evaluation of GBS

2.36 In the remainder of this chapter, we reflect on the issues arising for this evaluation from the perspectives on GBS, internationally and at country level, that we have developed during the inception phase.

2.37 All the study countries show changing approaches to budget support. Much of the change is incremental, suggesting scope for tracing the evolution of newer forms and for assessing the extent to which the pace of change, and the new forms adopted, reflected particular country contexts.

2.38 At the same time, strong international influences are at work: most notably, the evolution of PRSP approaches directly linked to the context of debt relief and the HIPC initiative in particular. One of the study countries (Uganda) was clearly ahead of the international trend towards partnership GBS (a pioneer). In other cases, the introduction of partnership GBS is at least partly an attempt to create an institutional context that in Uganda appeared as a precondition. This raises important issues for investigation, concerning the sequencing of reforms, the balance between demand and supply factors in the propagation of partnership GBS, and the consequent implications for ownership. The salience of HIPC as a motivation for governments to adopt PRSs raises questions about the degree of government ownership and the commitment involved, given the strong financial incentive to meet HIPC criteria.

2.39 Case-study countries show very different degrees of penetration of partnership GBS (in terms of the length of the partnership GBS history, the absolute and proportional volumes of aid involved, and the range of donors engaged). In at least three of the case-study countries (Vietnam, Nicaragua and Malawi) partnership GBS initiatives are so recent, or unconsummated, that there is no possibility at all of tracing any effects all the way to impact on poverty reduction. Here the appropriate focus for evaluation would seem to be on the relevance of moves towards partnership GBS, and what can be learned about entry conditions and interactions among stakeholders, by the early experiences of GBS. (Failures as well as successes can be instructive.) In some others (Mozambique, Uganda) partnership GBS dominates the aid landscape, at least in terms of aid management institutions and dialogue, though not necessarily overshadowing other aid modalities in financial terms. There is much more scope in such cases to take the evaluation to the level of outputs and outcomes, at least. Also the range of experience represented invites investigation of the extent to which the evolution of GBS reflects lesson learning within and across countries.

2.40 Compilation and inspection of the inventory also suggests patterns across countries in individual donors' characteristic approaches to partnership GBS. It will be important to understand the factors behind different donors' predispositions, and to check the extent to which they may be modified at country level through interaction with the preferences of governments (in different political and economic contexts) and through interaction with other donors (coalition formation and peer pressure are clearly significant factors in modifying donor approaches).

2.41 A further line of enquiry is suggested, in order to try to understand the different rates of adoption of partnership GBS across countries: to what extent do donors respond to partner country preferences in respect of the aid modalities that they adopt (including GBS) and more particularly, do donors apply the same criteria consistently across countries in their use of GBS and the conditions they apply to it?

2.42 Donors clearly do not all operate with identical perceptions about the objectives or the instrumentalities of GBS. It is obviously impractical to evaluate separately the financial GBS inputs of different donors (hence this joint evaluation), but GBS-contributing donors may nevertheless (a) have significantly different intentions and expectations in providing GBS, and (b) whether or not for that reason, also adopt significantly different designs of GBS in terms of the conditions, approach to dialogue, TA inputs and H&A concerns. At the least, this study should not assume that such differences are immaterial; most likely a careful examination of these differences of detail will yield significant lessons about the more and less effective ways of approaching partnership GBS.

2.43 The suggestion that disaggregation may be appropriate on the donor side, so as to explore differences among and interactions between different donors, has its counterpart on the government side. It is clearly inappropriate to regard governments as monolithic: indeed the analysis proposed by the Evaluation Framework centres on the way that GBS may influence incentives within governments and the wider political process. But there is a further point: it is too simplistic to view aid and donors as external influences on country systems. They clearly also operate as actors within those systems, but the degree to which this occurs or is made explicit varies enormously across the study countries. At one end of the spectrum, the government of Uganda invites its donor partners to participate directly in the budget formulation process (an explicit quid pro quo codified in a set of partnership principles), while, at the other, budget formulation processes in Vietnam are opaque even to many insiders. At both ends of this spectrum donors seek to influence both the short-term allocation of resources and the long-term

evolution of resource allocation systems. There appears considerable scope to learn lessons from the contrasting experiences of ostensibly similar instruments (e.g. PRSCs) across such different institutional environments.

2.44 Against this background, we turn in the next chapter to a consideration of the approach and methods by which such enquiries may most fruitfully be pursued.

3. Approach and Methods – Enhanced Evaluation Framework

Introduction

3.1 A key requirement for the Inception Report is to present an elaboration of the approach and methods for the study. The Evaluation Framework provides a foundation for the study, but the study team is required, in the light of its work during the inception phase, to propose improvements to the EF and incorporate them in a revised causality framework.²⁵ This chapter (a) notes the methodological challenges inherent in the evaluation of GBS programmes; (b) outlines the main features – and the strengths – of the Evaluation Framework; (c) describes the study team's approach during the inception phase (the systematic exploration of evaluation issues in the study countries while checking the ability of the EF to address them); (d) describes the Enhanced Evaluation Framework (EEF) that has been developed through this process, and how it addresses certain limitations of the EF; and (e) addresses some additional issues in methodology.

Challenges in GBS Evaluation

Complexity

- 3.2 The evaluation of GBS programmes is exceptionally complex, in a number of ways:
- (a) The initial inputs are themselves complex – a combination of funds with associated dialogue and conditionality, technical assistance and capacity building, harmonisation and alignment.
 - (b) Most of the initial inputs are not discrete (the GBS funds may be clearly and separately identified, but the other inputs are frequently bundled with non-GBS inputs).
 - (c) The desired ultimate effects are complex (poverty reduction in a number of dimensions across diverse countries).
 - (d) Changes in outcome and impact indicators will be partly (and sometimes dominantly) the effects of other causes (deliberate effects of non-GBS inputs, or exogenous factors).
 - (e) The chain of causality is a long one, both conceptually and temporally. Following a results chain all the way from inputs to impact is known to be challenging, particularly in moving from outputs to outcomes and impact (see the discussion of the "attribution gap" in ¶3.6 below). In any circumstances, the intervals between inputs and their immediate effects and outputs, outcomes and impacts will be significant. When effects are expected to result from processes of institutional change, the plausible interval for effects to be manifested is longer still. Moreover, results may be such that they are measurable only periodically and with difficulty; this lengthens the interval, in practice, before results can be ascertained. In some cases, moreover, confidence in the reliability of a link from hypothetical causes to observed effects may require repeated observations and evidence that the effect is persistent.
 - (f) In the case of GBS, many of the intermediate effects postulated are not in themselves straightforward to measure, let alone to attribute proportionately to multiple causes particularly in a dynamic context where GBS is only one of the influences on systems that are continually changing.

²⁵ The TOR refer to a "causality tree" but MG comments on the first draft of this IR acknowledge that "causality framework" is a more appropriate term. See ¶3.14 below.

- (g) The logic of causation is often itself controversial (for example, even if it could be demonstrated that GBS leads to the adoption of a particular policy designed to reduce poverty, the appropriateness and efficacy of the policy – either generally or in a particular country context – may well be disputed).
- (h) Last, but not least, the choice and the construction of appropriate counterfactuals (what would have happened if GBS had not happened?) is both difficult and controversial.

Basic Evaluation Principles

3.3 Complexity, in itself, does not make the basic principles of a sound evaluation any less relevant. Thus:

- The DAC evaluation criteria (Box 3.1) remain essential.
- It becomes even more important to distinguish as rigorously as possible, at all stages of the enquiry, between *findings* (facts), *conclusions* (interpretation of the facts, drawing on the judgement of the evaluators), and *recommendations* (reasoned advice based on the evaluation findings and conclusions).²⁶
- Assessment has to be made against an appropriate and explicitly identified counterfactual. This has both a conceptual dimension (what is the relevant alternative to the with-programme situation that the evaluators should consider?) and a practical one (is it practically possible to reconstruct a plausible without-programme situation?).²⁷

Box 3.1: The DAC Evaluation Criteria²⁸

The five DAC evaluation criteria are:

- *Effectiveness*: The extent to which the development intervention's objectives were achieved, or are expected to be achieved, taking into account their relative importance.
- *Efficiency*: A measure of how economically resources/inputs (funds, expertise, time, etc.) are converted to results.
- *Relevance*: The extent to which the objectives of a development intervention are consistent with beneficiaries' requirements, country needs, global priorities and partners' and donors' policies.
- *Impacts*: Positive and negative, primary and secondary long-term effects produced by a development intervention, directly or indirectly, intended or unintended.
- *Sustainability*: The continuation of benefits from a development intervention after major development assistance has been completed. The probability of continued long-term benefits. The resilience to risk of the net benefit flows over time.

²⁶ This is reflected in the reporting structure adopted (see Chapter 4 and Annex J).

²⁷ Appropriate counterfactuals are considered further in the final section of this chapter.

²⁸ As appended to the GBS study TOR (all taken from *Glossary of Key Terms in Evaluation and Results Based Management*. OECD-DAC (DAC Working Party on Aid Evaluation), Paris 2002).

*Causality and Attribution*²⁹

3.4 When causality is complex, the characteristic challenge for an evaluator is that of *attribution*:

- Can a particular observed effect be attributed to a particular observed cause?
- If so, to what extent? (Is it a major or minor cause? Is it a sufficient or a necessary cause? and so forth.)
- What degree of confidence in the attribution is justified?

3.5 The standard approach requires a careful construction of the logic of the programme (this may be variously termed a logical framework, a causality tree, a results chain, a logic chart, etc.). The logic chart spells out what a (project or) programme is trying to achieve, and enables the logic to be systematically (and consistently) tested. Crucially, by spelling out the links in the chain of results, this approach allows the different links to be separately tested.

3.6 GTZ have noted in the context of projects and programmes that an *attribution gap* exists: *Up to the level of 'use of outputs' attribution is relatively easy in most cases. However as we climb the levels of 'outcomes' and 'impacts' external factors that cannot be influenced by projects and programmes become increasingly important. The attribution gap widens up to an extent where the observed changes cannot be directly related to project inputs any more.*³⁰

3.7 The attribution gap is arguably more problematic for an intervention such as General Budget Support than it is for projects. Whereas with projects the immediate effect of project activities can often be readily discerned, with GBS the immediate effects may not always be clear since the intended effects lie at a system-wide level. Thus the problem of attribution becomes a serious one at an earlier level of the intervention logic. One response to the attribution gap is to monitor only up to outcome level. In the case of GBS that is unsatisfactory because attribution problems already exist at that level and the goal of GBS is quite clearly at a level beyond that. Rather, the approach must be to recognise that attribution is an issue at all levels, to avoid the trap of equating temporal sequence with causality, to use careful triangulation techniques, and always explicitly to consider the possibility that results observed may be attributable to factors outside the GBS programme.

3.8 Two prime requirements are thus:

- To set out clearly the logic that is being tested, why the particular hypotheses that are embodied in this logic are being tested, the types of evidence that are appropriate in testing them, and the degree of confidence with which particular attributions may be made.
- To be as transparent as possible about the process by which the evaluators proceed from findings to conclusions and (eventually) recommendations.

²⁹ See Mayne, J. (1999). *Addressing Attribution Through Contribution Analysis: Using Performance Measures Sensibly*. Ottawa: Officer of the Auditor General of Canada (Discussion Paper). Attribution analysis and causality analysis are not (as implied in some of the comments on IR2) alternative techniques but two sides of the same coin.

³⁰ GTZ (2004) *Results Based Monitoring: Guidelines for Technical Cooperation Projects and Programmes*.

3.9 A third requirement is to optimise the learning potential from the evaluation by identifying and focusing on a manageable number of main lines of enquiry. The scale and the complexity of GBS programmes mean that the number of possible causal chains is indefinitely large. The evaluation must select a sub-set for close examination based on the concerns of stakeholders, the evaluability of particular sub-chains, and the potential to add significantly to what is already known. At the same time, it must be clear how particular sub-chains fit into the overall logic of the programme, so that they contribute systematically to the overall assessment, both within and across the sample countries. This issue is taken up in Chapter 4, but we first proceed with the construction of a robust overall framework for the evaluation.

Strengths and Limitations of the Evaluation Framework³¹

Origins and Outline of the Evaluation Framework

3.10 The Evaluation Framework was commissioned on behalf of the OECD DAC Evaluation Network. It draws on an earlier GBS Evaluability Study³² produced for the Evaluation Department of UK DFID. The expectations, priorities and objectives of budget support have been unpacked and discussed in the EF which is:

... intended as a practical tool that can be used to guide a number of country-level joint evaluations. The ultimate purpose of these exercises is to assess whether GBS is a relevant, efficient, effective and sustainable mechanism for poverty reduction.

... 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 links in Logical-Framework fashion. (EF ¶S2, and §2)

3.11 The logical framework approach is not new to this field³³ but the Evaluation Framework is an elaborate and rigorous version, based on a very specific set of hypotheses about how GBS is meant to work. Annex C provides a comprehensive summary, and the full Evaluation Framework takes this to an impressive level of detail. Key features are:

- The standard logical sequence of five Levels (Inputs, Immediate Effects,³⁴ Outputs, Outcomes and Impacts) as depicted in Figure 3.1.
- The identification of two main sets of effects: flow-of-funds effects and institutional effects.
- Provision of detailed guidelines for research questions and approaches at each level of the framework, based on assessing whether 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 (though these are to be explicitly considered in asking why/why not questions related to attribution).
- The EF is designed for country-level evaluations, though, by providing a standard methodology, it is intended to facilitate a series of comparable country case-studies.

³¹ Booth, D. and Lawson, A. (2004). *Evaluation Framework for General Budget Support*. London: ODI.

³² Lawson, A., Booth, D., Harding, A., Hoole, D. and Naschold, F. (2002). *General Budget Support Evaluability Study, Phase 1: Final Synthesis Report*. Oxford and London: Oxford Policy Management and ODI.

³³ See White, H. (1999) *Dollars, Dialogue and Development: An Evaluation of Swedish Programme Aid*. Stockholm: Sida.

³⁴ "Activities" is a common alternative designation for this level.

TOR Requirements for Refinement of the EF

3.12 The TOR (§5.3) make clear that the EF is an essential platform for the present evaluation, but that it is not to be used uncritically:

The Framework is more general and broader in scope than the specific focus of this evaluation. Hence, it should be used as the basis and logical structure to the key themes and issues of the evaluation and to the proposed approach and method, but requires further details to become specific to the country case studies and this evaluation.

During the inception phase, the consultants shall, firstly assess the Framework in relation to the types and approaches of GBS in the different case-study countries and their objectives and conditions, and suggest any changes and/or additions to the Framework.

Secondly, with a focus on GBS (identified types, approaches and objectives), the evaluation team should break the Framework down into a causality tree. The tree should show the links between the different inputs and the results on the different levels and also the links and hierarchy between the different results and expected effects that are currently presented at the same level. Possible gaps and important inter-linkages and/or interdependencies should be highlighted and analysed and used in identifying the key themes and issues of the evaluation and when refining the approach and method. Furthermore, it is important to make explicit the intended as well as unintended positive and negative effects of GBS including major areas of risk and how they relate into the causality tree.

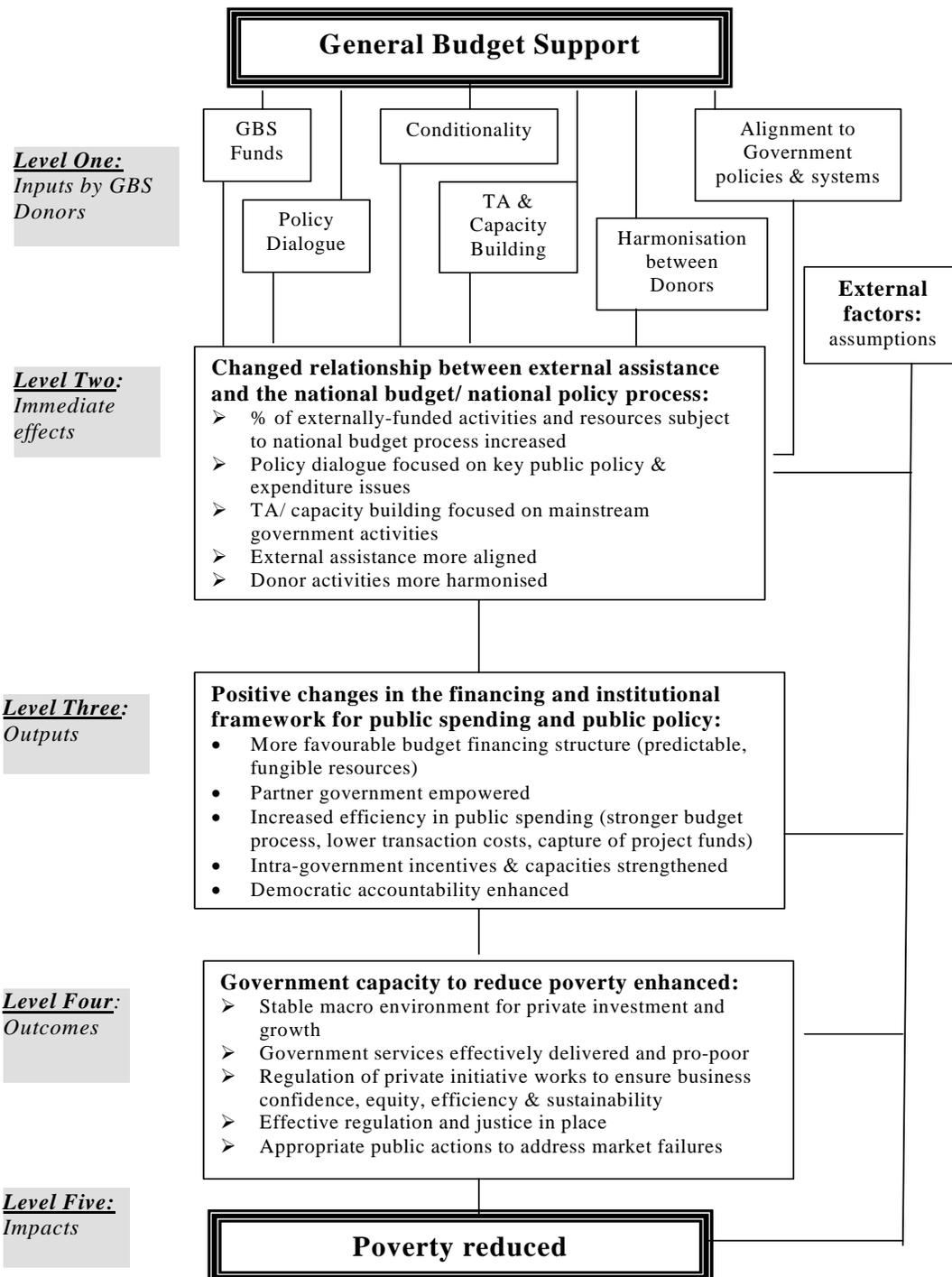
A single causality tree should be developed during the inception phase. This tree will be applicable to all country studies to allow cross-country lessons in the thematic and synthesis reports. However, different parts of the tree may be more or less relevant (but still considered) for the different case-study countries and the links may be more or less strong.

3.13 In their comments on the first draft of this inception report the Management Group describe the task thus:

A further elaboration of the approach and methods, including a comprehensive causality tree to be used as a framework (as hypotheses to be tested) during the remainder of the study. Making a ‘causality tree’ – causality framework is possibly a better name – includes:

- (i) Unpacking the GBS inputs,*
- (ii) Identifying the expected result chains from the inputs to the levels 2 to 5 from the EF, and*
- (iii) Defining the evidence to be looked for in order to evaluate whether the intended effects, outputs, outcomes and impact did materialise. (Management Group (2005) ¶1.8.)*

Figure 3.1: GBS Evaluation Framework (simplified version)



Source: Booth & Lawson 2004.

3.14 We agree that causality framework is a better name, because there are in practice two different, though complementary requirements. The first is actually to take a broader view than the EF does of what elements are relevant to be included within the evaluation.³⁵ The second (informed by the first) is to spell out, in causality tree fashion, particular, and more detailed results chains that merit special attention, and define the relevant evidence to be looked for accordingly. It thus makes sense to insist (a) on a single overall causality framework to guide all the country evaluations, and (b) that the location of each of the results chains being investigated is clearly mapped onto the overall framework; but (c) it is not practical to depict all relevant sub-chains *in detail* on the same diagram or matrix.

The IDD Approach

3.15 The approach adopted by the IDD team has sought to build on the impressive strengths of the Evaluation Framework, while enhancing it with a number of linked analyses. This approach was outlined in the IDD proposal and has been elaborated in the course of inception work. The IDD team has benefited from, and greatly appreciates, the close collaboration of the ODI team (including the original authors of the Evaluation Framework), which has been undertaking a separate evaluation of GBS in Tanzania based on the EF.³⁶

3.16 The IDD team³⁷ has been organised into interlocking country teams and an overall synthesis team which reflects five main analytical perspectives: partnership analysis, macroeconomic analysis, analysis of public finance management (PFM), institutional analysis, and poverty analysis. The five analyses were selected because each provides a different perspective on the relevance, effectiveness, efficiency, impact and sustainability of GBS. Triangulation among these analyses, as well as triangulation of indicators within them, is intended as a means throughout the study to check findings and to develop a more complex and rounded understanding of the influence of GBS on governmental systems, financial flows, institutional realities, service delivery and poverty. These analytical perspectives are in no sense an alternative to the Evaluation Framework; rather they are a systematic approach (a) to using it and (b) to verifying its utility.

3.17 These five perspectives have been complemented by systematic attention to the four cross-cutting issues (CCIs) identified in the TOR – democracy & human rights, gender, HIV/AIDS and the environment: see Annex H for full details on the approach to, and conclusions from, the review of CCIs at the inception stage. Thus the approach throughout the inception phase has been to use the EF as a platform, and to use the five analyses as perspective, simultaneously checking what the key issues are, and how well the EF works as a means of evaluating them.

³⁵ The TOR reference to the EF being "more general and broader in scope than the present evaluation" is best understood as referring to the EF's tendency towards operating as a comparative evaluation of competing aid modalities; in other respects, as discussed below, it may actually adopt too narrow an interpretation of the relevant factors to be incorporated within the evaluation.

³⁶ Booth, D., Lawson, A., Williamson, T., Wangwe, S. and Msuya, M.: *Joint Evaluation of General Budget Support Tanzania 1995–2004. [2004a] Inception Report; [2004b] Phase 2 Report. Preliminary Assessment of Efficiency & Effectiveness of Budget Support and recommendations for improvements; [2004c] Final Report.* Reports to the Government of Tanzania and to the Poverty Reduction Budget Support (PRBS) Development Partners. Dar es Salaam and London: Daima Associates Limited and ODI.

³⁷ See Annex B for details of the study team and their topics and countries of focus.

3.18 Inception visits to the case-study countries gathered details of GBS and related programmes (see the inventory of GBS at Annex E), introduced the study to key contacts in government, donor offices and civil society through workshops and individual meetings, collected essential documents (see Annex D for general and specific bibliographical material) and identified the main developments in government–donor relationships in GBS over the last decade. Progress reports and interim country reports (though not formal deliverables of the evaluation) analysed initial information in terms of the five analyses and evaluation framework, recorded initial ideas regarding important causal relations in GBS programmes (as contributions to the causality framework), and suggested priorities for the fieldwork phase.

3.19 In the process of preparing this inception report, we further explored causality relations from the perspectives of the five analyses, using knowledge gained from the inception visits. This influenced our perception of how the EF needs to be further developed, and refined our judgment of the appropriate focus of some of the five analyses (and CCIs). The two results of this process are the Enhanced Evaluation Framework (EEF) that is explained in the next section, and the key evaluation questions (Chapter 4).

The Enhanced Evaluation Framework

Assessment of the Evaluation Framework

3.20 Like any logic chart, the Evaluation Framework is a considered simplification. Simplification is necessary in order to make the task of evaluation more manageable, but there is also a risk of obscuring, or assuming away, elements that it would be better to keep in view. The more debatable simplifications in the Evaluation Framework are as follows:

- (a) It does not systematically address entry conditions. By what criteria is a country deemed to be (and remain) eligible for GBS? In practice there is a great deal of debate as to what are and should be such criteria, both for initial provision of GBS and for its possible interruption or termination. What are the different contexts in which GBS may be adopted, and how do they influence the design of GBS?
- (b) Although it unbundles the GBS inputs (into funds, dialogue, conditionality, TA, harmonisation and alignment), it does not take into account that the non-financial inputs are commonly themselves bundled with non-GBS inputs and activities. (For example: harmonisation and alignment activities related to GBS are often part and parcel of broader H&A efforts; significant TA that supports GBS objectives is often provided through project modalities and rather tenuously linked to the GBS funds.)
- (c) It treats GBS donors as a homogeneous group (almost as a single actor). (This is thus one example where it ignores interactions within a level.) Moreover, its language implies that there are GBS-donors and other donors, whereas in practice the GBS donors themselves also provide aid through other modalities.
- (d) It is oriented towards a comparison (and contrast) of GBS with other modalities, but does not systematically explore the interactions between GBS and these modalities. (These can work both ways: the effectiveness of GBS may depend on complementary projects, e.g. for TA and capacity building; at the same time, improvements in policies or in institutions that are attributable to GBS may also make non-GBS aid more effective.)
- (e) In keeping with the previous simplifications, it generally treats "new GBS" as a single design. Moreover, it imposes the evaluators' normative logic (what the objectives of GBS should be), and the fact that some of the donors may see the logic differently – both from the evaluators and from each other – is discarded. It might be argued that

this is consistent with the joint donor approach to evaluation: donors accept that it is not appropriate or practical to attribute the results of GBS separately to individual donors; nevertheless (i) their different approaches may influence different designs of GBS, which are a legitimate concern for evaluators, and (ii) their different expectations may affect how they perceive and react to the performance of GBS, and hence its sustainability as an approach.

- (f) This last point relates to the issues of feedback and circularity. The Evaluation Framework essentially portrays the logic as one-directional – from inputs to impact. However, GBS is part of policy and budgetary systems (both the government's and the donors') that are characteristically circular: successive inputs are influenced by feedback from earlier inputs. There is an additional circularity within the logic of GBS itself: improvements in many of the factors that are treated as minimum requirements for GBS to be feasible (e.g. a government's basic fiduciary standards) are themselves regarded as part of what GBS can accomplish (its outputs and outcomes).
- (g) It is much stronger on the "PFM-focused" aspects of the logic (discretion to formulate and manage budgets, etc.), and on the immediate flow-of-funds effects, than on the less "mechanical" effects; for example, the causality chains for the expected changes in policies that influence growth are not spelled out.
- (h) A consequence of these and other simplifications is that many key influences on the performance of GBS are treated as assumptions and not well-specified.

3.21 The EF is also rather weak on the time-scale for effects. There are important leads and lags throughout: at what interval should an input of GBS have an impact on the volume of service delivery? on the quality of service delivery? on income poverty? GBS is meant to stimulate institutional change, but this takes time and requires learning, and so forth. This is an issue that has important design implications for GBS. For example, if donors seek a design that rewards a government's performance in using GBS, it is rather important to take account of the time intervals between elements of the results chain (see earlier discussion, ¶3.2(e)).

Overview of the Enhanced Evaluation Framework

3.22 An Enhanced Evaluation Framework (EEF) is schematically presented in Figure 3.2. In it we address the shortcomings of the original Evaluation Framework as follows:

- (a) A new "Level 0" is introduced so that design context and entry conditions can be systematically addressed:
 - From a donor perspective, it is important to consider factors that make a partner government eligible for GBS.
 - It is also important, though, to consider the criteria that make a donor willing to commence (and to persist with) a GBS programme.
 - Where minimum standards are concerned, they may provide a benchmark for later performance evaluation.
 - It is important to consider feedback to the donor, and to the donor's constituents, since this will affect the continuity and durability of GBS as a modality.
- (b) There is more recognition in Level 1 (inputs) of parallel inputs, both from donors and from government. It is not intended to expand the scope of evaluation to cover all aid or all government inputs, but there are several reasons to make these inputs more visible:
 - Some are impossible to unbundle (as already noted).

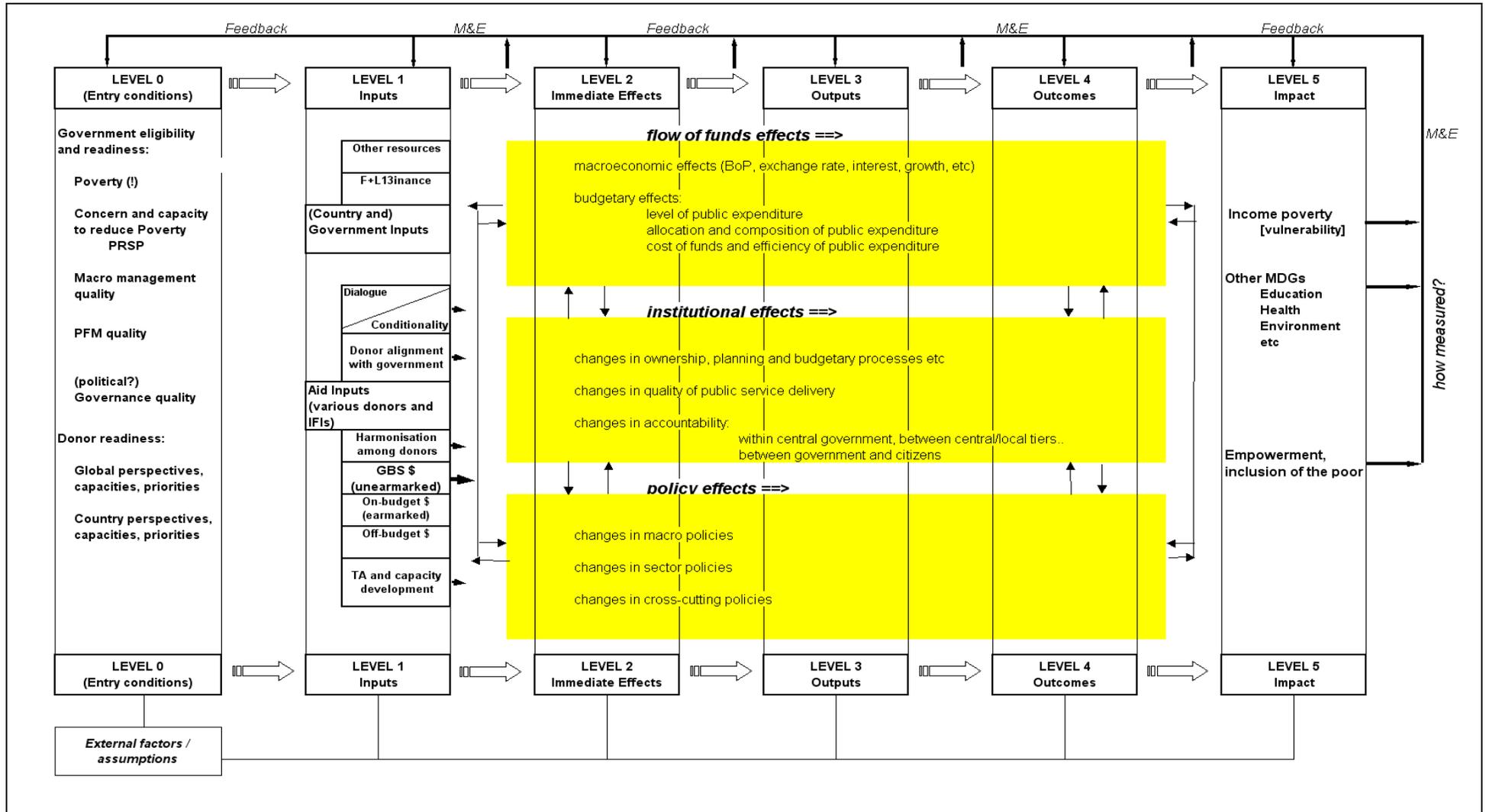
- Just as it is impractical to separate the effects of different donor inputs, so it may be necessary first to consider the combined effects of GBS and government inputs, before proceeding to attribute (a proportion of) those effects to GBS.
 - It highlights the importance of considering the interactions (positive as well as negative) between modalities.
 - It depicts the close link between dialogue and conditionality (it is not helpful to consider formal conditionality in isolation).
- (c) The effects from Levels 2 through 4 (Immediate Effects/Activities, Outputs, Outcomes) are conceived as three streams, not just two (funds, institutions, and also policies). It is stressed that these are not seen as separate compartments: as depicted in the diagram, there are systematic interactions between funds, policies and institutions. However, explicit inclusion of policy as a causal mechanism helps resolve some of the difficulties in the EF (for example, that "public actions to address market failures" appear at Level 4 of the EF without any real explanation of the intervening transmission mechanism).
- (d) The different poverty dimensions at Level 5 (impact) are unpacked. This recognises that different causal chains may influence some of the different dimensions. Notably, public expenditures may have a direct impact on education, health and other dimensions where government services can play a direct role, while income poverty is less susceptible to such direct effects, and the drivers of empowerment are as much political as economic. The EEF presentation highlights the fact that the different dimensions need to be separately considered.
- (e) Feedback loops (from Level 5 and intervening levels) are depicted. This is consistent with the earlier observation that GBS programmes are characteristically iterative. Special attention should be paid to the systems for M&E at each level. (The "how measured?" question applies at all levels of the framework.) Such monitoring is of course an important source of information from which to assess the effects of GBS. More immediately, what is being monitored, by whom, and how, are factors that have a direct bearing on the relevance and sustainability of the design of GBS programmes. Although it is impossible to include a time scale within the diagram (because different intervals apply to different components), systematic attention to the feedback loops will bring this consideration to the fore.

3.23 There is still a parallel set of "external factors/assumptions" as in the EF, but a lot has been taken out of it and made more explicit.

Using the EEF

3.24 This Enhanced Evaluation Framework answers to the broader parts of the TOR specifications (¶3.12–3.14 above). It addresses the systemic weaknesses in the Evaluation Framework that were our starting point, and it provides a more comprehensive framework, common to all the study countries, for showing how key sub-chains of the evaluation relate to each other. It does not discard the specific hypotheses that were embodied in the original EF, but allows them too to be set in a broader context, to be more rigorously posed and tested, and to be supplemented by additional hypotheses at different levels of detail. This task is carried forward in the identification of key evaluation questions in the next chapter, and also in Annex G, where we present a comprehensive revision of the detailed Evaluation Framework (corresponding to, and supplanting, the EF's *Chapter 6: The Evaluation Framework in Detail*).

Figure 3.2: The Enhanced Evaluation Framework (schematic view)



Additional Methodological Issues

Standard Terminology

3.25 Inception work has highlighted the lack of standard definitions and common understandings surrounding GBS (not only the definitions of GBS and other forms of programme aid but also those of a wide range of related terms that are employed in the discussion). This clouds the debate, with much talking at cross-purposes. Moreover, a standard approach across the study countries requires care that the IDD study teams are all using the same definitions.

3.26 Annex F (Terminology) provides a common reference point for the study teams. Wherever possible we adhere to standard DAC terminology, but there are important areas that this does not cover. We expect that further refinement of definitions and classifications in certain areas (e.g. conditionality, predictability) will be required as we take the evaluation to further levels of detail.

Counterfactuals

3.27 Any evaluation requires consideration of the counterfactual, and the evaluation will be meaningful only if the counterfactual is a relevant one. 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.

3.28 For aggregate flow of funds and budgetary effects, it is certainly appropriate to ask if GBS is additional or a substitute for other forms of aid, and frame the counterfactual accordingly. (GBS may be a substitution now but considered as a possible addition in future, in which case both alternatives could be considered.) Of course, deciding what is the appropriate counterfactual, in principle, does not mean necessarily that it is practical to model one in econometric detail. For budgetary effects we should ask – even if we decide we cannot answer – what is the marginal effect of GBS on public expenditure, taking fungibility into account? At the same time, it is important always to consider whether, for the purposes under investigation, GBS is materially different from other forms of aid. (For example, it is not immediately clear, when Dutch disease is at issue, that the effects of more GBS should be considered any differently than the effects of the same amount of aid in a different form.) It is important not to stray towards the Herculean task of evaluating aid as such.

3.29 For assessing relevance and appropriate design of GBS, we have to take account of donor intentions and rationales: if GBS is put forward as a corrective to certain deficiencies in other forms of aid (high transaction costs, say), then a relevant counterfactual is persistence with those forms of aid, and we have to ask both whether the original diagnosis was correct (previous modality did have the characteristics identified) and whether GBS performs better in the relevant dimensions. In doing so we will pay attention, as already noted, to interactions, both positive and negative, between different forms of aid.³⁸

³⁸ This resonates with the March 2005 Paris Declaration, endorsed by all member organisations of the SG:

We acknowledge that enhancing the effectiveness of aid is feasible and necessary across all aid modalities. In determining the most effective modalities of aid delivery, we will be guided by development strategies and priorities established by partner countries. Individually and collectively, we will choose and design appropriate and complementary modalities so as to maximise their combined effectiveness.

3.30 There are also aspects of GBS where the appropriate comparator will be alternative designs of GBS itself. Even here, though, we have already noted that many of the dimensions of GBS are common to GBS and other modalities. As the TOR observe:

The shift to GBS has also resulted in increased attention to key issues of development co-operation such as ownership, partnership, transaction costs, coordination and alignment, which make an evaluation of GBS highly relevant to the development cooperation context in general. (TOR §2.3)

It is highly likely that GBS-specific findings related to such aspects will also be relevant to the choice and design of a wider range of aid instruments.³⁹

3.31 This approach to counterfactuals is reflected in the guidance on key evaluation questions, which are the central concern of the next chapter.

³⁹ See Annex F (Terminology) for more on the distinction between aid modalities and aid instruments, and for a classification of the most relevant dimensions of aid instrument design.

4. Key Evaluation Questions

Introduction

4.1 This chapter sets out the main evaluation questions to be addressed through the Enhanced Evaluation Framework. The Revised Inception Report noted the need for further refinement and prioritisation of evaluation questions prior to field work. The present chapter has been thoroughly re-drafted in the light of further work prior to and during the Field Preparation Workshop (26–27 April 2005). It deals with:

- Causality hypotheses and the detailed evaluation questions at each level.
- The prioritisation of enquiries within the evaluation.
- Principal causality chains and a causality map.
- Key evaluation questions.
- The structure for country reports.
- Cross-cutting issues and themes.
- The limitations of the evaluation.

Main Causality Hypotheses and Detailed Evaluation Questions

4.2 Box 4.1 shows, for each level of the logical framework, the main effects that are hypothesised to result from GBS. These draw on the hypotheses depicted in the original Evaluation Framework, but have been elaborated to reflect the wider considerations introduced by the Enhanced Evaluation Framework (see Chapter 3 above). These hypothesised effects form the first column (the "logical sequence") of the detailed evaluation questions which are presented in Annex G.

4.3 As well as the logical sequence depicted in Box 4.1, Annex G sets out:

- the general questions appropriate to each level of the EEF (these are linked to the OECD DAC evaluation criteria);
- specific questions (and relevant indicators) related to each of the numbered links in the logical sequence;
- the principal assumptions to be checked by the evaluators at each level.

4.4 Annex G thus corresponds to Chapter 6 of the original EF ("The Evaluation Framework in Detail"). It has been thoroughly revised to reflect the causality hypotheses in this chapter, to sharpen and streamline the associated evaluation questions and indicators, and to incorporate the evaluation questions suggested by the review of cross-cutting issues (Annex H). Annex G nevertheless remains a work in progress (as is normal in such cases – the equivalent section of the original EF makes the same point).

Box 4.1: Enhanced Evaluation Framework – Logical Sequence of Effects

Level 1 (the design)
<p>1. Adequate quantity and quality of inputs are provided by new GBS:</p> <p>1.1 Funds</p> <p>1.2 Policy dialogue</p> <p>1.3 Conditionality</p> <p>1.4 TA/capacity building linked to</p> <ul style="list-style-type: none"> • PFM • Pro-poor sectoral policies and good governance <p>1.5 Alignment and harmonisation</p> <ul style="list-style-type: none"> • IPs' alignment to government goals and system • IPs' harmonisation
Level 2 (the immediate effects/activities)
<p>2.1 More external resources for the government budget (additionality)</p> <p>2.2 Proportion of external funds subject to national budget process increased (increased fungibility)</p> <p>2.3 Increase in predictability of external funding of national budget</p> <p>2.4 Policy dialogue and conditionalities focused on pro-poor policy framework and improved PFM</p> <p>2.5 TA/capacity building established to:</p> <ul style="list-style-type: none"> • improve PFM processes including budgeting, accounting, financial control, audit • improve the linkage between PFM and pro-poor sectoral policies and good governance <p>2.6 Actions to ensure IPs' alignment are in place</p> <p>Actions and agreements to improve IPs' harmonisation are in place</p>
Level 3 (the outputs)
<p>3.1 Increased resources for service delivery:</p> <ul style="list-style-type: none"> • External resources are treated as additional • Cost of funding budget deficit reduced <p>3.2 Partner government is encouraged and empowered to strengthen PFM and government systems:</p> <ul style="list-style-type: none"> • To use the budget to bring public sector programmes into line with government goals, systems and cycles (PRSP/MTEF) • To set up performance monitoring systems to measure the effectiveness of public expenditure at the level of the final beneficiaries • To promote alignment and harmonisation by IPs <p>3.3 Partner government is encouraged and empowered to strengthen pro-poor policies:</p> <ul style="list-style-type: none"> • To establish and execute an adequate sequence of reforms to ensure macro-economic stability and private sector development • To establish and execute pro-poor policies and targeting in health, education, agricultural and rural development • To enhance social inclusion policies, through decentralisation and participation of the civil society, reform of the administration of justice and respect for human rights <p>3.4 Improved aggregate fiscal discipline:</p> <ul style="list-style-type: none"> • More predictable funding flows • Incidence of liquidity shortfalls reduced, hence less use of Central Bank overdrafts and less accumulation of arrears <p>3.5 Operational efficiency of public expenditure is enhanced:</p> <ul style="list-style-type: none"> • By reductions in certain types of transaction costs to partner government (e.g., non-standard procurement systems, brain-drain effects of parallel project management structures) • Better planning, execution and oversight reduces wasteful spending, controls corruption better, spreads positive lessons across the public sector

3.6 Allocative efficiency of public expenditure is enhanced:

- By a more effective budget process: multi-year, results oriented, transparent, participatory; with effective execution and audit; with an adequate tracking system
- By increased capture of project funds in budget
- By stakeholders taking the domestic budget more seriously (because that's where the money is)

3.7 Intra-government incentives and capacities are strengthened:

- Official reporting lines are more respected (vertical through government to cabinet, not horizontal to IPs)
- Public-service performance incentives are strengthened, so that policies are made and implemented, audit and procurement systems work, and corruption is reduced

3.8 Democratic accountability is enhanced:

- Greater role of parliament in monitoring budget results
- Accountability through domestic institutions for IP-financed spending is enhanced
- Conditions for all-round democratisation are thereby improved, including the trust of people in their government and hence their level of expectations

Level 4 (the outcomes)

4.1 Macroeconomic environment is favourable to private investment and growth:

- Inflation controlled
- Realistic exchange rate attained
- Fiscal deficit and level of domestic borrowing sustainable and not crowding out private investment

4.2 Regulation of private initiative works to ensure business confidence, equity, efficiency and sustainability:

- Policies on corruption, property rights resolutely pursued
- Market-friendly institutions developed

4.3 More resources flowing to service delivery agencies

4.4 Appropriate sector policies include public actions to address major market failures, including those arising from gender inequalities

4.5 More effective and accountable government improves administration of justice and respect for human rights, as well as general confidence of people in government

4.6 More conducive growth enhancing environment

4.7 Public services effectively delivered and pro-poor:

- Service delivery targets met for key pro-poor services
- Evidence of increased use of services by poor (including poor women)

Level 5 (the impact)

5.1 Income poverty reduction

5.2 Non-income poverty reduction

5.3 Empowerment and social inclusion of poor people

Focusing and Prioritising the Evaluation

4.5 Box 4.1 follows from the main causal hypotheses proposed or implied by advocates of GBS. It is clear that these are potentially very broad in scope, and give rise to a very large number of possible lines of enquiry when detailed causal mechanisms (causality sub-chains) are taken into account. It is neither practical nor appropriate to investigate all the possible sub-chains with the same intensity.

4.6 In prioritising among possible enquiries within the overall evaluation it is appropriate to consider:

- *Significance of the issue* (potential value-added from the evaluation): how interested are stakeholders in the issue, and can the lessons learned potentially make a difference to the effectiveness of future GBS?
- *Susceptibility to evaluation*: this may relate to methodological issues, to the availability of data (and cost of collecting it), and also to whether there is a sufficient history of GBS (altogether or in a specific country) to support an evaluation to the level implied.
- The available *resources and time-scale* of the study.

4.7 The precise focus and emphasis of the evaluation may differ between study countries, according to their different circumstances and interests (indeed, as discussed in Chapter 2, this is essential given the range of GBS experience and the particularly short history of partnership GBS in some of the study countries), but all will be related to the common structure of the EEF. The rest of this chapter sets out a common causality map and detailed evaluation questions which are common to all the study countries and which are tightly linked to the structure for the Country Reports.

Principal Causality Chains and Causality Map

4.8 The Revised Inception Report noted:

Particular issues to be investigated will correspond to causality sub-chains within the EEF, and such sub-chains may function both within and across levels. In each case, we will identify the levels involved and locate the sub-chain on the overall "map" that is provided by the schematic view of the EEF in Figure 3.2 (previous chapter).

4.9 This is done in Figure 4.1, which maps the causality hypotheses implicit in Box 4.1 (and Annex G) onto the schematic EEF of Figure 3.2. The arrows do not show every possible causal link but are used to highlight what we judge to be the principal ones for investigation. In turn, the key evaluation questions (see below) are cross-referenced to particular causality sub-chains depicted in Figure 4.1. We must stress that this identification of principal sub-chains is preliminary: it may well be that other sub-chains, or particular variants within a sub-chain, will emerge as particularly important in some or all of the study countries, while some sub-chains that are prominent in the rationale for GBS will turn out not to be significant in practice. Figure 4.1 is a guide to enquiry, not (yet) a presentation of results.

Key Evaluation Questions

4.10 Box 4.2 summarises the nine key evaluation questions (EQs), which will structure the field work and the country reports. These EQs have been deliberately framed so as to (a) follow an appropriate sequence through the levels of the EEF, (b) cover all areas of the EEF as depicted in Figure 3.2, with a minimum of overlap, and (c) to focus on the principal causality sub-chains depicted in Figure 4.1. Annex K, which complements Annex G, provides much fuller details: it sets out judgement criteria, relevant evidence and sources of data for each EQ. It also suggests the appropriate counterfactuals for each EQ, consistent with the approach to counterfactuals described in Chapter 3.

Structure for Country Reports

4.11 The structure for reports needs to strike a balance between the rigour imposed by any logical framework and the natural boundaries of the topics under discussion. As noted, this has been taken into account in framing the EQs. Annex J sets out the detailed report structure for the Country Reports, and this is summarised in Box 4.3.

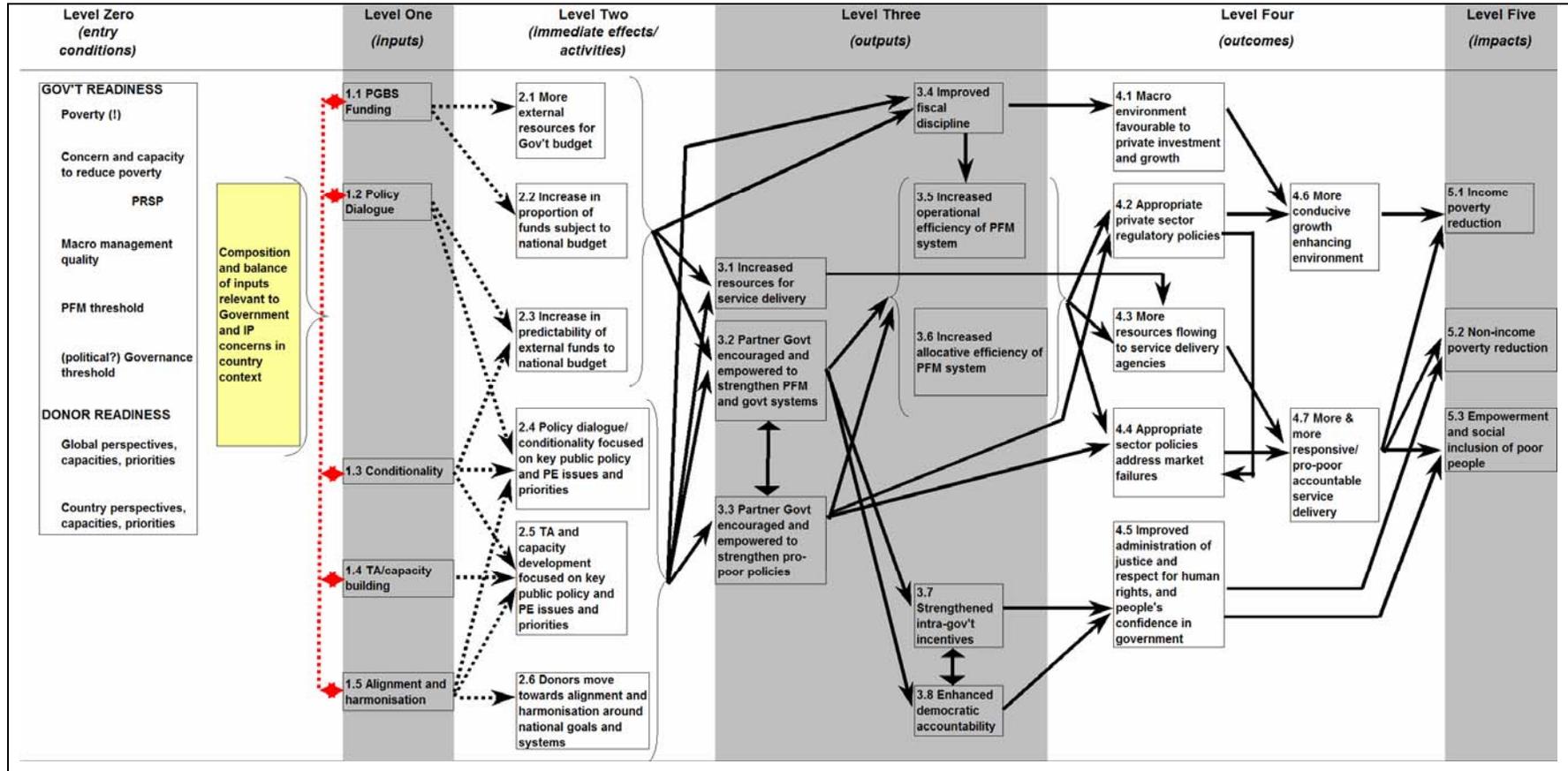
4.12 Part A of each Country Report provides the context for partnership GBS, including identifying relevant GBS and related programmes and describing their evolution in the light of the attitudes and expectations of national and international partners (Level Zero of the EEF).

4.13 Part B is the analytical section (there will be no recommendations in this Part). The first nine chapters in Part B each respond to one of the key Evaluation Questions summarised in Box 4.2. In each case the chapter will (briefly):

- (a) Relate the scope of the chapter to the EEF (which levels and streams of the EEF is it *mainly* concerned with?).
- (b) Note the main causal hypothesis (hypotheses) that is (are) being tested (cf. the causality sub-chains identified in Box 4.2).
- (c) Note any special challenges in attribution associated with the hypothesis(es).
- (d) Present relevant findings (facts).
- (e) Draw overall conclusions based on the judgement criteria of the EQ matrix (Annex K). In doing so, evaluators will indicate both.
 - The strength of the apparent causal link from partnership GBS (PGBS) to a particular effect (no effect, weak, moderate, strong);
 - The evaluators' confidence in the attribution (high confidence, medium confidence, low confidence).

4.14 A tenth analytical chapter (not based on a separate EQ) requires teams to draw together findings relating to a number of cross-cutting issues and themes, based primarily on the analysis performed for each of the EQs (see the next section for discussion of the principal cross-cutting issues and themes identified).

Figure 4.1: Causality Map for the Enhanced Evaluation Framework



Box 4.2: Key Evaluation Questions

Evaluation Questions	Levels / DAC Criteria	Principal Causality Chains
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?	Levels: 1←0 <i>Relevance</i>	(Relevance question from Level 0 to Level 1; considerations of internal consistency)
2. Has PGBS contributed to greater harmonisation and alignment of the aid process?	Levels: 2←1 <i>Effectiveness and efficiency</i>	1.5 (and other inputs) → 2.6
3. How efficient, effective and sustainable has been the contribution of PGBS to the performance of the public expenditure process?	Levels: 3←1 (flow of funds) <i>Effectiveness, efficiency and sustainability</i>	2.2/2.3→3.2→3.5/3.6 2.4/2.5→3.1
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?	Levels: 3←1 (institutional effects) <i>Effectiveness, efficiency and sustainability</i>	2.4/2.5/2.6→3.2→3.5/3.6/3.7/3.8
5. How efficient, effective and sustainable has been the contribution of PGBS to improving public policy processes and policies?	Levels: 3←1 (policy flow) <i>Effectiveness, efficiency and sustainability</i>	2.4/2.5/2.6→3.3→3.5/3.6
6. How efficient, effective and sustainable has been the contribution of PGBS to macroeconomic performance?	Levels: 4←1 (flow of funds) <i>Effectiveness, efficiency, sustainability</i>	2.1/2.2/2.3→3.4→4.1→4.6 2.4/2.5/2.6→3.4→4.1→4.6
7. How efficient, effective and sustainable has been the contribution of PGBS to improving government performance in public service delivery?	Levels: 4←1 (institutional effects) <i>Effectiveness, efficiency and sustainability</i>	3.5/3.6→4.4→4.7 3.3→4.4→4.7 3.1→4.3→4.7
8. How far has PGBS strengthened government impact on poverty?	Levels: 5←1 <i>Impact and sustainability</i>	3.2→3.7→4.5→5.2/5.3 3.2→3.8→4.5→5.2/5.3 4.6→5.1 4.7→5.3
9. Is the PGBS process itself sustainable?	Levels 5←0 (feedback loops) <i>Sustainability</i>	Feedback loops as illustrated in EEF diagram (Figure 3.2)

Source: Extracted from Annex K – the full matrix of key Evaluation Questions, including judgement criteria, evidence, data sources, counterfactuals.

Box 4.3: Chapter Structure of Country Reports

Executive Summary

Part A: Context/Description

- A1. Introduction and Conceptual Framework for the Evaluation
- A2. The Context for Budget Support in [Country]
- A3. The Evolution of Partnership GBS in [Country]

Part B: Evaluation Questions: Analysis and Main Findings

- B1. The Relevance of Partnership GBS
- B2. The Effects of Partnership GBS on Harmonisation and Alignment
- B3. The Effects of Partnership GBS on Public Expenditures
- B4. The Effects of Partnership GBS on Planning and Budgeting Systems
- B5. The Effects of Partnership GBS on Policies and Policy Processes
- B6. The Effects of Partnership GBS on Macroeconomic Performance
- B7. The Effects of Partnership GBS on the Delivery of Public Services
- B8. The Effects of Partnership GBS on Poverty Reduction
- B9. The Sustainability of Partnership GBS and its Effects
- B10. Major Cross-Cutting Issues and Themes

Part C: Synthesis – Overall Conclusions and Recommendations

- C1. Synthesis of Evaluation Conclusions

Annexes

Source: Annex J.

4.15 Part C of each Country Report (CR) will develop an overall assessment of partnership GBS in the case-study country, built up from (and systematically cross-referenced to) the component assessments in Part B. The focus will be on which approaches to GBS appear to have worked/not worked at country level and why: it is the Synthesis stage which will address the *"in what country contexts"* element of the TOR. This part will include preliminary recommendations at country level, together with tentative suggestions towards the cross-country synthesis. The draft CRs are, in the first instance, an input to the Synthesis Note workshop (see Chapter 5 for full explanation of the study sequence). Country level recommendations will be finalised only after the SN workshop and the SG review of the SN and CRs.

4.16 Annexes will be used to present detailed supporting evidence, where necessary, in order to keep the main text as concise as possible.

4.17 It is anticipated that the Synthesis Report will follow a similar structure, but this will be subject to review at the Synthesis Note workshop (see Chapter 5).

Cross-Cutting Issues and Themes

4.18 The TOR identifies four cross-cutting issues (CCIs) – gender, environment, HIV/AIDS, and democracy & human rights. The approach to these CCIs within the evaluation is fully described in Annex H, and they will be considered, as appropriate, under each of the EQs. The final chapter in the analytical section of the Country Reports will be used to draw together, and if necessary amplify, findings on these CCIs.

4.19 The same chapter will also be used to provide a consolidated treatment and additional analysis of a number of other major issues and themes which cut across several EQs, and/or which have been highlighted as especially significant. These include:

- *Decentralisation* (in particular how decentralisation affects and is affected by different designs of GBS).
- *Government capacity and capacity building* (considered both as a prerequisite for and an objective of partnership GBS). This links to the issue of absorptive capacity and implications for GBS as a modality for the proposed scaling up of aid.
- *The balance between public and private sectors* and how this may be influenced by GBS.
- The effect of GBS on *transaction costs* in aid management. (It is important to take an overall perspective on transaction costs, including the relevant transaction costs of expenditure and implementation at levels 2 and 3, and not only the negotiation and management transaction costs at levels 0 and 1.)
- *Corruption* and its implications for the sustainability of GBS and the risks attached to it.
- *Ownership* (including a disaggregation of ownership issues – e.g. who on the government side is the owner and what does this imply for relationships and effectiveness of partnership GBS?).
- *Democratic accountability* and whether/when accountability to donors detracts from or reinforces the strengthening of institutions for democratic accountability.

Limitations of the Evaluation

4.20 It is important to keep in mind the following inherent limitations of this evaluation:

- While it is important for evaluators to validate whether there has been an increase in poverty-related budgeting, and to review its effectiveness at local levels, it is neither practical nor appropriate to undertake original analyses of the determinants of poverty, and we also have to rely for the most part on secondary sources for our analysis of poverty levels and trends, and indeed many other relevant indicators.
- The evaluation is not about the effectiveness of aid per se. It is about whether GBS is an effective way of delivering aid (in which contexts and with which designs).
- As already noted, given the time-scale (the short period during which partnership GBS has been in place, even in the countries where it is well established), the study will clearly be able to say rather more about relevance and design than about effectiveness and impact.
- Taking the preceding points together, in considering the relevance of GBS and particular approaches to it, the study will draw on standard theories (about drivers of poverty reduction, about effective PFM, and so forth), e.g. the theories embodied in DAC guidelines and other standard sources to which (by and large) the GBS donors subscribe.
- This links to a general point: it is appropriate for the study to focus largely on second order issues (e.g. how do CCIs enter the dialogue? how is poverty monitored? are private sector issues considered in dialogue, in policy analysis, in monitoring? and so on) since it is at this level that it will be possible to consider whether the dynamics of GBS address such issues directly and indirectly.

4.21 We are mindful that this is not intended to be a comparative evaluation of GBS against other aid modalities, and we have already indicated that such a stark question is in any case at odds with the subtleties and dynamics of the aid relationships that we observe. However, it is important also to check that the evidential standards by which GBS is evaluated are not substantially more (or less) demanding than the standards that it is possible to apply to the evaluation of other forms of aid.

4.22 A final note is in order on the limitations of the type of evaluation framework that forms the core of the methodology for this study. As one of the internal quality assurance reviewers for this study has noted:

The wish to impose a formal structure could be taken too far:

- a. For reasons set out in the inception report, it is in the nature of GBS that it is supposed to have systemic effects on government activities and programmes which are subject to numerous other influences – perhaps unfortunately, but inevitably this means that attribution must always require an element of judgment which must take account of the macro setting, and that judgment must be in light of the particular national conditions.*
- b. Even if this were “narrow” economics, attributing cause and effect of such a general intervention would be very difficult. However, at its very core, GBS is about political economy – the key intention (and question) relates to shifts in political behaviour (government decision-making quite broadly defined) as a result of a change of mode of aid delivery. This involves interpretation of political/administrative processes; the use of a log-frame or causality tree approach may prove to be too mechanical to pick up the subtleties of institutional change.*

[It is important therefore] not to push the effort for a formal, more or less rigid, framework too far, which might lead [the study] to miss the complexities that result from the diversity of institutional arrangements and political economy in the countries being studied.

5. Work Plan

Introduction

5.1 Work during the inception phase (described in ¶3.15–3.19 of Chapter 3) has provided a strong platform for the remainder of the study. A wealth of country-level and more general material has already been collected. This includes, but is not limited to, what appears in the (informal) country inception reports.⁴⁰ The present chapter explains the work plan – the process for producing the remaining outputs of the study in a way that adequately addresses the issues highlighted in Chapter 4 on the basis of the methodology described in Chapter 3. This chapter (a) presents the revised timetable for delivering the major outputs of the study; (b) explains the process aspects of ensuring the quality and comparability of the study's findings; and (c) discusses the remaining stages and outputs in more detail.

Overall Timetable

5.2 Figure 5.1 shows the revised timetable for the remainder of the study, as agreed by the study team with the Management Group. The stages following the submission of the revised inception report are summarised in Box 5.1 below. As shown in Figure 5.1, there is some overlap in the dates of these stages, and activities within stages are often concurrent.

Box 5.1: Main Study Phases and Activities, Post-Inception

Stage (dates)	Activities
Preparation for Field Studies (14 March – 30 April 2005)	<ul style="list-style-type: none"> • MG and SG review of revised Inception Report; final revision of IR in response to comments • Further preparation of detailed field study guidelines • Field study preparation workshop (26–27 April)
Country Field Studies (May – July 2005)	<ul style="list-style-type: none"> • 7 country field visits between May and July (see Annex I for details of proposed country timing and focus) • Submission of draft country reports
Field Study Synthesis and Overall Synthesis Report (July – November 2005)	<ul style="list-style-type: none"> • Review and quality assurance of draft Country Reports (CRs), and submission to MG (all submitted by 12 September) • Synthesis Note (SN) workshop (28 July) • Drafting of SN; submission to MG by 12 September • SG meeting on CRs and SN (week of 3 October)
Study Finalisation and Dissemination (December 2005 – March 2006)	<ul style="list-style-type: none"> • Finalisation of Country Reports (all submitted by 14 November) • Drafting of Synthesis Report (submission to MG by 21 November) • SG meeting on Synthesis Report (week of 12 December) • Final draft SR to MG (week of 9 January 2006) • Preparation of Note on Approach and Methods (submitted by 23 January 2006) • Dissemination Conference (14–15 March 2006)

⁴⁰ The first draft of this Inception Report included more detail of this country-level material. Subsequent drafts have focused more strictly on the study methodology and the main issues that emerge from the inception work.

Quality Assurance and Comparability

Internal Quality Assurance

5.3 The following quality assurance measures are in place:⁴¹

- (a) A core team of subject specialists, which interlocks with the country study teams, so that country work plans and drafts are systematically reviewed from all the different analytical perspectives represented in the core team, while all the core team members also participate in field work in at least one of the study countries.
- (b) An additional panel of specialists in the cross-cutting issues (democracy and human rights, gender, HIV/AIDS, and environment) has helped to develop the study methodology, and will also review the Synthesis Note and the Synthesis Report.
- (c) The core team includes a decentralisation specialist, who will participate in one of the field studies and advise on all aspects of decentralisation issues.
- (d) A two-person panel of distinguished experts reviews all substantive drafts before their formal submission to MG/SG. (In practice, the study also submits drafts for informal peer review to other experts, from within and outside the organisations that constitute the study consortium, at no cost to the client.)
- (e) The study is also supported by two evaluation specialists (one of whom also acts as a quality assurance reviewer of substantive drafts). One of these specialists focuses particularly on the evaluation methodology embodied in the EEF. The second will participate in, and facilitate, the field preparation workshop and the synthesis note workshop with particular attention to standardising approaches across countries.
- (f) The study team leader is committed virtually full-time to the study throughout its core period (from May through September 2005). A full-time research assistant provides additional support, and the study is also able to draw on corporate support from IDD and other members of the consortium.
- (g) The study team regards close consultation and cooperation with the Management Group, based on transparency, as a further important aspect of quality assurance.

Ensuring that the Country Studies can be Synthesised

5.4 The country studies are required to be high-quality reports in their own right and of practical value to partner governments and other stakeholders at country level. This means that they must take into account the particular context and concerns of each participating country. At the same time it must be possible to synthesise the country reports so as to draw valid conclusions that transcend individual country experiences. Measures to ensure both the quality and the comparability of the country reports include:

- (a) Application in all cases of the common EEF (Chapter 3, and detailed in Annex G).
- (b) The common set of key Evaluation Questions and the common report structure described in Chapter 4 and detailed in Annexes J and K.
- (c) The use of consistent terminology (see Annex F).
- (d) Workshops for the core team prior to field work to ensure that study guidelines and research instruments are standardised and understood the same way (see Box 5.2).
- (e) Particular attention in the field study workshop to the adoption of the standard methodology and the rigorous observance of good practice in evaluation (with reference to DAC and other guidelines).

⁴¹ See Annex B for full details of personnel and their assigned roles.

- (f) Reference to standard international benchmarks where these are relevant (including the Public Expenditure and Financial Accountability (PEFA) indicators for PFM, OECD DAC guidelines on donor harmonisation and alignment, and the summary governance indicators published by the World Bank). However, such benchmarking systems need to be used very judiciously: they should inform, but cannot substitute for, the evaluators' considered assessment.
- (g) A standard format for the country reports (see Annex J).
- (h) Testing the draft country report format by the preparing a "prototype" for Uganda.⁴²
- (i) Country team leaders will be required to prepare a "zero draft" (a country-specific detailed outline) of the country reports prior to commencement of the field visits. This will ensure a thorough stocktaking of what material is already available (both from secondary sources and from the initial field visits), and what may need to be followed up during the country visits (in addition to the particular additional areas of focus identified for each country). It will also enable the full draft country reports to be prepared more rapidly after the field visits.

Stakeholder Consultation

5.5 Consultation with appropriate stakeholders is an essential aspect of the study. Their views and experiences are not only an important part of what is to be researched, but an important guide in ensuring that the study focuses on lesson-learning in the areas that concern them. Formal consultations with the MG and SG are built into the study timetable already described. Additional consultation and interaction will occur at country level:

- Informally, through interviews and meetings as part of the research.
- More formally, through systematic liaison with the country reference groups that have been established.
- Through workshops during the field visits: these have to be guided by the wishes and availability of the country stakeholders (led by the government), but the field teams will aim, as a minimum, to provide a detailed briefing on the study so far at the beginning of the field visit, and to hold a more formal workshop (or workshops) towards the end of the field work. The latter workshops will be designed to allow the study team to present its work in progress, to highlight key issues or controversies, and to promote further detailed discussion by a range of stakeholders.⁴³
- Through the sharing of draft working papers and reports with the country reference groups. (The team notes that formal dissemination of all reports requires the prior approval of the MG, particularly if any preliminary recommendations are included, and expects that SG members will liaise with their country-level staff to ensure that appropriate protocols are followed in the distribution and citation of any drafts.)

Field Studies and the Country Reports

5.6 The key steps involved in the field studies and Country Report (CR) preparation are shown in Box 5.2.⁴⁴

⁴² It was not possible to prepare a complete prototype before the field preparation workshop, but work on it revealed potential problems in the detailed EEF and in the IR2 country report structure, and fed directly into the elaboration of the EEF, the report structure and the EQs now reflected in Annexes G, J and K respectively.

⁴³ As noted in Annex H, the CCIs will feature systematically in such workshops.

⁴⁴ CT = Country Team; CTL = Country Team Leader; TL= [overall study] Team Leader; QA = Quality Assurance.

Box 5.2: Main Steps in Field Studies and Country Report Preparation

Steps	By	Comment
Preparation for Field Work		
<ul style="list-style-type: none"> • Scheduling of country visits 	CTLs, liaising with country contacts	Country preferences are decisive with regards to the timing of visits. Tentative schedules shown in Annex I to be confirmed once IR is accepted.
<ul style="list-style-type: none"> • Finalising field study guidelines <ul style="list-style-type: none"> – Final revision of IR – Refinement of Annex G (EEF in detail) – Refinement of CR outline 	TL	Upon receipt of MG/SG comments (IR itself is an important manual for the study teams).
	QA and Synthesis Team	Further refinement/elaboration, prior to and at the Field Preparation Workshop.
	TL [and Uganda team]	Outline to be tested with Uganda "zero draft" and refined on the basis of experience and workshop comments.
<ul style="list-style-type: none"> • Field preparation workshop <ul style="list-style-type: none"> – Plan workshop – Participate in workshop 	TL, Synthesis Team and QA	<i>This draws on the experiences of the initial workshop (13-14 September 2004), and the inception workshop (23-24 February 2005).</i>
	Synthesis Team and CTLs	Planning team to include evaluation specialists and workshop facilitator.
<ul style="list-style-type: none"> • Additional central support <ul style="list-style-type: none"> – Work on PEFA indicators – Web site 	TL and Synthesis Team [plus] [PFM team, liaising with CTs]	Collate basic minimum sub-set of PFM indicators across all study countries [in liaison with PEFA].
	[Web site manager]	Reorganise web site, and place on it copies of (or links to) all key documents, at both country and synthesis levels.
<ul style="list-style-type: none"> • Additional preparation by country teams <ul style="list-style-type: none"> – Review country-level secondary material – Prepare detailed country outline 	CTLs	Linked, inter alia, to cross-country studies identified in Annex D, and to development of internal web site.
	CTLs	For review by TL prior to field visit.
Field Work		
<ul style="list-style-type: none"> • Country Research • In-country consultation 	CTs	Special country focus indicated in Annex I
	CTs	In-country workshops, etc.
Country Report Drafting & Review		
<ul style="list-style-type: none"> • Report drafting <ul style="list-style-type: none"> – Full draft CR – Revised draft CR 	CTL [with CT]	To be submitted to TL within 3 weeks from completion of field work.
	CTL	After MG/SG comments and after Synthesis Note stage (see Box 5.3).
<ul style="list-style-type: none"> • Report review <ul style="list-style-type: none"> – Informal circulation to country reference group – Internal team review – Quality assurance – MG and SG 	CTL to circulate	(At or prior to submission of full draft) following guidelines agreed with MG and country reference group.
	Synthesis Team	Formally, on submission of full draft (ST also to provide informal support during drafting).
	QA	Comments on full draft.
	MG	See schedule in Figure 5.1.

The Synthesis Note

5.7 The Synthesis Note (SN) is an important bridge between the Country Reports and the Synthesis Report. As described in the TOR (§6, §5.4.3):

The Synthesis Note shall summarise the main findings of the field missions. The note should address all key issues and themes agreed in the inception phase. The note should be succinct and mainly constitute a basis for a SG meeting to prepare the synthesis phase. The note will not be published as a self-standing report.

This note will ... allow for discussions and for a first validation of the preliminary findings before the drafting of the final synthesis report starts.

5.8 Steps concerning the Synthesis Note are shown in Box 5.3. The SN workshop, scheduled for 28 July, is currently planned as an internal workshop for the core study team. It will take place in Birmingham. It is intended to review the draft country reports and discuss the issues to be highlighted in the SN, as well as those to be further developed in the final Synthesis Report. It will also provide feedback on the draft CRs that will be reflected in their finalisation. The study team considers that there would be considerable added value in expanding the peer review element of the SN workshop by inviting selected additional specialists (e.g. academics, aid agency practitioners), although this is clearly beyond the resources of the study budget itself. Participation by the experts retained to advise the SG during the study would also be welcomed.

Box 5.3: The Synthesis Note Stage

Steps	By	Comment
<ul style="list-style-type: none"> • Country Reports 	CTLs	See Box 5.2. Some of the field visits will not be completed before July, and the final date for submission of all CRs to MG is 12 September. CTLs for the later studies will ensure that at least an Executive Summary of Main Findings is available for the SN workshop.
<ul style="list-style-type: none"> • Synthesis Note Workshop <ul style="list-style-type: none"> – Preparation – Participation 	TL and QA (evaluation specialist) Core team [Synthesis Team and CTLs]	To prepare and facilitate effective workshop that extracts main lessons from country research and prepares basis for Synthesis Report while also refining and standardising the CRs.
<ul style="list-style-type: none"> • Synthesis Note Drafting and Review <ul style="list-style-type: none"> – Draft SN – Internal QA – Review by MG 	TL and Synthesis Team QA and CCI MG	Internal draft for QA prior to submission to MG. SN draft to be reviewed by Cross-Cutting specialists as well as internal QA. To be submitted to MG, along with all CRs by 12 September.
<ul style="list-style-type: none"> • SG Meeting on Country Reports <ul style="list-style-type: none"> – Presentation of SN and CRs 	TL and selected core team members	Week commencing 3 October 2005. SN as key focus for discussion of field findings and SR issues and requirements.

The Synthesis Report

5.9 The Synthesis Report (SR) is the most important output of the study. It is intended to be more than a summary of the Country Studies and to provide *global lessons and recommendations to donors and partner Governments at both operational and policy levels* (TOR §5.1). It is further described as follows (TOR §6):

The Synthesis Report should be based on the country reports and the thematic reports as well as other relevant documents including the literature review. It is important that the data, findings, conclusions, lessons and recommendations of the country evaluations and the thematic evaluations are effectively used when drafting the synthesis report. The synthesis report should include references to the other reports and clearly account for the evidence on which it is based. Furthermore, the synthesis report should not only summarise the findings of the country evaluations and the thematic evaluations but should add another level of analysis to draw out the more general trends, findings, conclusions, lessons and recommendations about the future use of GBS, both operationally and at the policy level.

5.10 It does not now appear that thematic studies formally linked to the main GBS evaluation will have taken place by the time the SR is prepared.⁴⁵ Nevertheless, there is a great deal of relevant recent and ongoing work, under the auspices of the DAC, SPA and other bodies, by individual agencies, academics and NGOs, that it would be appropriate to take into account. There is scope too for checking the experiences of the seven case-study countries against some others where GBS or similar activities have been the subject of available reports, studies and evaluations. Box 5.4 shows the main steps for preparation and review of the Synthesis Report.

Box 5.4: The Synthesis Report Stage

Steps	By	Comment
<ul style="list-style-type: none"> • Synthesis Note and Workshop • SG Meeting 		<p>See Box 5.3: this will be the platform from which preparation of the SR will start.</p> <p>Week of 3 October – to review SN and CRs and provide further guidance on the SR.</p>
<ul style="list-style-type: none"> • Drafting and Review of SR <ul style="list-style-type: none"> – Preparation of draft SR – Review of draft SR – SG meeting on SR – Response to MG/SG comments and final draft 	<p>TL With Synthesis Team</p> <p>CTLs, CCI specialists and QA</p> <p>Presentation by TL and selected team members</p> <p>TL</p>	<p>Thorough internal review before submission to MG (21 November 2005).</p> <p>Week of 12 December 2005.</p> <p>Week of 9 January 2006.</p>
<ul style="list-style-type: none"> • Dissemination Conference 	MG/SG	14/15 March 2006.

⁴⁵ A parallel SECO study to gather preliminary lessons learned on the best international practice of GBS PAFs is taking place, and its preliminary and final outputs are being shared with the GBS study teams.

The Note on Approach and Methods

- 5.11 The final output to be delivered by the study team is a *Note on Approach and Methods*:
A note discussing the approaches and methods used in the evaluation, and in particular the experiences of using the Evaluation Framework, should be developed at the completion of the evaluation. This note should provide lessons for future evaluations of GBS and feed into the development of guidelines for evaluation of GBS. (TOR §6)
- 5.12 This note will be prepared by the study Team Leader, in consultation with other members of the study team.