



WHAT WE KNOW ABOUT THE EFFECTIVENESS OF BUDGET SUPPORT

Evaluation Synthesis

2017



DEval

GERMAN
INSTITUTE FOR
DEVELOPMENT
EVALUATION

For a number of years, budget support has been a preferred aid modality for implementing the principles of effective aid formulated in the 2005 Paris Declaration. Various evaluations find that aid provided through this modality contributes to increased budget allocation towards poverty reduction and improved development outcomes in a number of countries. Nonetheless, budget support has increasingly come under criticism in recent years, and many bilateral donors have either partly or fully stopped using this modality.

Against this background, DEval conducted a comprehensive and systematic review of existing evidence on the effectiveness of budget support, based on a considerably larger number and greater variety of sources than previous synthetic work. To close important remaining knowledge gaps this evaluation synthesis draws together different perspectives on the impact of budget support from a total of 95 evaluations and studies.

The findings substantiate existing knowledge on the effectiveness of budget support and at the same time generate new insights for decision makers in German and international development cooperation.

WHAT WE KNOW ABOUT THE EFFECTIVENESS OF BUDGET SUPPORT

Evaluation Synthesis

2017

Imprint

Published by

German Institute for
Development Evaluation (DEval)
Fritz-Schäffer-Straße 26
53113 Bonn, Germany

Tel: +49 (0)228 33 69 07-0

E-mail: info@DEval.org

www.deval.org

Authors

Magdalena Orth
Johannes Schmitt
Franziska Krisch
Stefan Oltsch

Responsible

Dr Stefan Leiderer

Design

MedienMélange: Kommunikation!
www.medienmelange.de

Editing

Jannet King
Dr Marcus Klein

Photo credits

nuu_jeed/Shutterstock (Cover), Patrick Lienin/photocase.de (Chap. 1), SasinTipchai/Shutterstock (Chap. 2), NUK2013/Shutterstock (Chap. 3), Bongkarn Thanyakij/Alamy Stock Foto (Chap. 4), criene/photocase.de (Chap. 5), connel/Shutterstock (Chap. 6), Dirk Hinz/photocase.de (Chap. 7)

Bibliographical reference

Orth, M., J. Schmitt, F. Krisch, S. Oltsch (2017), *What we know about the effectiveness of budget support. Evaluation Synthesis*, German Institute for Development Evaluation (DEval), Bonn.

Printing

Druckerei Bonifacius,
Paderborn



© German Institute for Development Evaluation (DEval)
As at: May 2017

ISBN 978-3-96126-049-2 (Print)

ISBN 978-3-96126-050-8 (PDF)

The German Institute for Development Evaluation (DEval) is mandated by the German Federal Ministry for Economic Cooperation and Development (BMZ) to independently analyse and assess German development interventions.

Evaluation reports contribute to the transparency of development results and provide policy-makers with evidence and lessons learned, based on which they can shape and improve their development policies.

This report can be downloaded as a PDF file from the DEval website:
www.deval.org/en/evaluation-reports

Requests for print copies of this report should be sent to
info@DEdeval.org

Acknowledgements

We extend special thanks to the members of the reference group, who provided highly valuable inputs and comments during the preparation of this evaluation synthesis. Evaluation, thematic and regional units of the German Federal Ministry for Economic Cooperation and Development (BMZ), the KfW Development Bank, the GIZ (Gesellschaft für Internationale Zusammenarbeit), as well as the German Development Institute (DIE) were represented in the reference group. For the quality assessment of budget support evaluations, the team received important feedback and support from international budget support experts.

We owe our gratitude to our peer reviewers Dr Martin Noltze and Dr Nadia Molenaers, who provided valuable and thorough comments for quality assurance.

Internally, we are grateful for Kristen Schubert's and Teresa Vogel's essential contribution to a smooth evaluation process. Special thanks goes to the interns Anne van der Kolk, Thomas Kockmeyer, Susanna Bolz and Simon Müller and student assistant Theresa Müller. They all supported the team with outstanding research assistance and inputs during major phases of the evaluation synthesis.

We cordially thank all contributors and supporters for the time and energy they invested in the evaluation synthesis on budget support.

EXECUTIVE SUMMARY

Background

Around the turn of the millennium, international development cooperation underwent a paradigm shift. Mounting evidence showed that over the past decades development aid had often failed to achieve its frequently shifting, yet always ambitious, goals. The evidence also suggested that the ineffectiveness had at least partly to do with the way aid was provided.

Past approaches to development cooperation were based on tightly donor-controlled projects and political conditionality of often limited credibility. In many cases they failed to produce the intended development outcomes. Instead, donor-driven and often highly fragmented aid projects lacked systemic effects on the social, economic and political development in aid-recipient countries, among other reasons caused by unintended effects. Often, this form of development cooperation, although aimed at strengthening domestic institutions and governments, instead established additional layers of aid-management institutions, that increased transaction costs, and undermined democratic control, ownership, and local capacities.^{6 i}

Besides identified shortcomings in development projects, criticism has also been directed at the so-called structural adjustment programmes (SAP) of the Bretton-Woods-Institutions. With the SAP, International Monetary Fund (IMF) and the World Bank provided concessional loans to indebted or bankrupt countries, linked to the recipient government's commitment to implement a predefined reform plan. In retrospect, these SAPs did not achieve the desired objective to reshape economic governance in the recipient countries.⁷

Based on the problematic experience with conventional project aid and SAP, donors and recipients were committed to rearranging aid relations and to increasing the effectiveness of development cooperation. In the course of several summits, the international community formulated step-by-step a new agenda for more effective aid. In 2005, with the Paris Declaration on Aid Effectiveness, the agenda was summarized in five central principles for more effective aid:⁸

- ownership by recipient governments
- the alignment of donor support with recipient strategies and systems
- the harmonization of procedures and approaches among donors
- results-orientation
- mutual accountability

To implement these principles in practice, new aid modalities had to be adopted, commonly subsumed under the term programme-based approaches (PBA). PBAs require ownership by the partner government, a formalized process for donor co-ordination, harmonization of donor procedures, and stronger use of local systems.

Belonging to the PBAs, budget support became increasingly popular towards the second half of the 2000s as one of the most consequent modalities for implementing the principles of effective aid formulated in the 2005 Paris Declaration. Usually provided jointly by multiple donors, its core objectives are to support the implementation of a country's national strategy for poverty reduction (financing function) and to promote good governance through support for jointly agreed reform processes (governance function). While general budget support (GBS) represents a non-earmarked contribution to support government spending, in sector budget support (SBS) the funds are usually earmarked for utilisation in a specific sector.

The budget support package consists of financial inputs, channelled into the treasury of the recipient country and of non-financial inputs, which support the reform process of the recipient. The non-financial inputs include conditionalities, a policy dialogue and technical assistance/capacity development (TA/CD). Budget support conditionality consists of two pillars: underlying principles and a Performance Assessment Framework (PAF). Underlying principles form the conditionality basis for providing general budget support. PAFs are negotiated among donors and partner governments and consist of performance indicators that help to monitor the government performance in the reform process. Disbursements of budget support instalments are linked to performance in the PAF. The policy dialogue between donors and partner governments

ⁱ In the entire evaluation synthesis, sources are listed in endnotes (Arabic numeral system) while explanations or illustrations are given in footnotes (Latin numeral system).

takes place on a technical level, in particular to assess the partner government's performance in PAF indicators, and on a strategic, higher-ranking level. TA/CD, as third non-financial input, serves to strengthen the capacities of the partner government and institutions to increase the effectiveness and efficiency in the use of the financial input. The combination of financial and non-financial inputs is expected to have reciprocal effects, increasing the effectiveness of both financial and non-financial inputs in their impact to reduce poverty.

Soon, a controversial debate on budget support started among development experts. In this debate two views dominate. Among the proponents, budget support, in particular GBS, is seen as particularly suitable instrument to intensify ownership of the partner governments, alignment with national structures of the partner and harmonization among donors. In the combination of financial inputs and the support of sector-specific and cross-sectoral reforms, systemic impacts on the partner country are expected. The reduction of poverty is to be achieved indirectly in supporting the activities of the partner government and the use of national systems of the partner. In contrast, the debate also revolves around fiduciary and political risks of budget support in case the quality of the partner governments' systems is insufficient or changes to the negative. As fiduciary risks, the debate discusses the misappropriation of budget support funds, such as corruption, as well as a misallocation of funds in non-poverty relevant sectors. On the side of political risks, budget support opponents fear negative effects on domestic accountability due to the one-sided support of the executive, and less incentives to mobilize domestic revenues.

During early stages of budget support programmes, the prime objective was to finance recipient countries' poverty-reduction strategies. Later, mostly among bilateral donors, budget support's function to promote good governance shifted to the centre of interest.⁹

The growing use of budget support led to an increase in evaluation work, analysing the effectiveness through an intervention logic recognized by the OECD Development Assistance Committee (OECD DAC).¹¹ The intervention logic

describes the expected effects of budget support. The achievement of those effects is supported by implementing the principles of effective aid, such as alignment to recipient governments' policies and systems, harmonization among donors and increased predictability in the disbursement of funds. By using the systems of the recipient country, also a reduction in transaction costs is expected. Expected effects include increased government spending in social sectors, strengthened PFM, less corruption, and improvements in domestic accountability and democratic governance. On outcome and impact level, the expected effects are improved macroeconomic performance and reduced income poverty, as well as an increase in service delivery of public goods and reduced non-income poverty.

Some of these joint donor evaluations point to positive effects of budget support regarding the financing function, and indicate impacts in the area of public financial management (PFM). In spite of these evaluations and their findings, they have had hardly any influence on the critical debate about budget support. The evaluations were perceived as single cases, taking the perspective of individual donors or organizations. To remedy this perception, evidence from evaluations has been summarized in a number of synthesis studies. However, the synthesis studies are often based on a limited number of cases and it is not always possible to identify which of the presented effects are widely backed by evidence.

In recent years, the acceptance of budget support in donor countries has declined significantly, causing many donor governments to suspend or stop budget support. Meanwhile, on the background of a vast financing gap for achieving the Sustainable Development Goals (SDG), the debate on suitable financing approaches mostly focuses on new finance modalities. The discussion covers a broad spectrum of modalities, but not all of them can be described as new. Prominently discussed are results-oriented modalities. The so-called results-based aid defines measurable outputs and outcomes, and financial means are disbursed once the agreed output or outcome is achieved. The aim is to incentivize partner governments to achieve the desired outputs or

¹¹ The intervention logic is described in the "Comprehensive Evaluation Framework (CEF)", developed as part of the methodological approach for evaluating budget support under the aegis of the OECD DAC and the European Commission (EC).

outcomes and to reduce fiduciary risks. While some modalities link all payments to the achievement of predefined results, others only incorporate results-oriented components. This is the case for budget support, using variable performance tranches to link single disbursement tranches to the success in defined PAF indicators. The similarity of the modalities in their performance orientation, the focus on outputs and outcomes and the use of recipient countries' systems allows to derive lessons from the case of budget support for results-based aid - even more so as practical experience with newer forms of results-based aid, such as Cash on Delivery, is so far rare.¹⁰

Objective

Against this background, this evaluation synthesis addresses the task of systematically reviewing existing evidence related to the effectiveness of budget support. With the intent to close important knowledge gaps in previous work, the synthesis was based on a large number of sources (95) and analysed different perspectives on the impact of budget support. Also, the evaluation synthesis analyses the qualitative heterogeneity of individual evaluations in a quality assessment and consistently describes the coverage of evidence for all findings.

The aim of this evaluation synthesis is to provide lessons learned for the design and implementation of future budget support programmes and related financing instruments. Through the systematic review of the evidence the aim is also to contribute to an objective debate on the impact of budget support.

To this aim, the evaluation synthesis answers the following two evaluation questions:

1. Which effects of budget support are substantiated by reliable evidence?
2. Under which contextual conditions does budget support generate results?

Methodology

The present evaluation synthesis adopts a theory-based approach, analysing the expected outcomes described in an intervention logic of budget support, recognized by the OECD

DAC. The intervention logic is also used in many similar evaluations. The evaluation synthesis includes a total of 95 different sources: 32 evaluations, 42 academic papers, and 21 “grey” literature sources. The sources cover analyses on GBS and SBS. The evaluation synthesis thus offers broad coverage and the analysis of different perspectives on budget support effects.

The results of predominantly qualitative evaluations and studies are systematically synthesized in line with the standards of a “systematic review”. Consequently, this evaluation synthesis follows a three-step process that includes an explicit search strategy, clear inclusion criteria, and systematic and software based coding and analysis. As central element of the methodology, a quality assessment of the sources was carried out. The quality assessment shows high homogeneity in the methodical quality of the evaluations, as well as scientific and “grey” literature. The overall methodological quality is sufficiently high to use all reports as the basis for this evaluation synthesis.

Aiming to assess the reliable effects of budget support, a distinction was made between empirically proven and inadequately documented effects on the basis of the coverage in the sources. The effects of budget support were analysed on two levels: effects of the entire budget support programme and effects of individual budget support inputs (financial and non-financial inputs). Effects are only accepted as reliable in the evaluation synthesis when they are described by at least 10 different sources. Also, based on the quality assessment, the sources were classified into two categories, “best evidence” and “second-best evidence” based on their methodological quality. This classification has been used for a robustness test, which is positive if results are based on sources from the “best evidence” category or on sources from both categories.

Findings

The findings summarized below are based on both quantitative and qualitative synthesis results related to effects of budget support, as presented in the intervention logic.

In this summary of findings, there is a clear differentiation between findings for which there is strong evidence in the

sources (frequent and consistent occurrence in the sources), and which can thus be viewed as reliably established effects, and those for which no clear picture emerges due to insufficient coverage in the sources. An overview of all empirically established effects is presented in Table 9.

From this evaluation synthesis, the following main findings emerge:

- Most expected effects of budget support are backed by sufficient evidence.
- Most proven effects of budget support are positive or at least in presence of certain context factors positive (e.g. high number and quality of donors' common interests and the adherence of partner governments to the conditionalities).
- Observed effects are mostly attributed to the whole budget support programme. Information on the specific effects of individual inputs of budget support (financial input, conditionalities, policy dialogue and TA/CD) is rarely provided.
- A multitude of sources convincingly describes the attribution of observed changes to budget support on output and induced output level. On the outcome and especially the impact level, attribution of observed changes to budget support programmes is often not plausibly supported by evidence.
- Important and highly debated risks of budget support are inadequately researched by the covered sources, thus statements on risks are not possible.
- The strength of budget support effects depends strongly on context conditions, such as institutional capacity and political will for reform of the recipient government.

Aid-effectiveness principlesⁱⁱⁱ

Budget support programmes have a positive effect on alignment of programmes to strategies and approaches of the recipient, consequent harmonization of donor programmes, and ownership of the recipient countries, although this effectiveness depends on the number and quality of common interests among donors and between donors and partner governments, as well as the adherence of partner governments

to the conditionalities attached to budget support programmes.

Budget support programmes have a positive effect on the reduction of transaction costs, but not on predictability. This is due to the delay of aid disbursements and uncoordinated decisions by donors following breaches of the conditionalities by partner governments. Transaction costs increased after the introduction of budget support, but declined again after joint processes were established. When donors withdraw from budget support, the suspension of tranches increases transaction costs again, as only small sums are paid out in relation to the operating expenses.

The non-financial inputs of budget support, in particular policy dialogue and conditionality, have a positive effect on harmonization, but do not contribute to increase ownership. Conditionality might even have a negative effect on predictability.

Regarding other specific budget support inputs, generalizable conclusions on the effects on aid-effectiveness principles cannot be drawn due to insufficient evidence. For example, no statement can be made about the effect of specific inputs on the alignment of the donor programmes to the procedures of the partner governments and, in the case of the transaction costs, no valid conclusion can be drawn regarding non-financial inputs. For all five principles, the inputs' effect on TA/CD is insufficiently covered.

Government expenditure

Budget support, especially the financial input, increases public spending. This increase is mostly apparent in social sectors such as health and education. One of the suspected risks of budget support provision was that this aid modality would reduce incentives for governments to raise domestic revenues and, as a consequence, crowd out domestic revenue; the "crowding-out effect". However, the sources analysed in this evaluation synthesis do not offer evidence that access to budget support funds reduces the mobilization of domestic revenues. At the same time, there are no or only slight improvements in domestic revenue mobilization through PFM reforms, putting at risk the sustainability of budget support

ⁱⁱⁱ These principles are derived from the intervention logic of budget support and are not identical with the Principles of the Paris Declaration, but comprise the alignment of donor support with recipient priorities and systems, the harmonization of procedures and approaches among donors, the ownership and responsibility by recipient governments, predictable disbursements of aid and a reduction in transaction costs.

effects on (pro-poor) public spending.

While effects of financial budget support inputs are often mentioned, the sources offer insufficient evidence on the effect of non-financial inputs on government expenditure. Due to this substantial deficit, no generalizable conclusions on the expected positive influence of non-financial inputs going beyond the sole financing function of budget support can be drawn for future programmes or related aid instruments.

Public financial management

Budget support programmes have a positive effect on the PFM of recipient countries, especially on budget formulation and planning, and the comprehensiveness and transparency of the budget. The improvements in PFM are best covered by the evidence. These improvements are specifically attributed to the non-financial inputs policy dialogue and TA/CD, as priority is generally given to improving PFM within these inputs. In particular, more recent evaluations recognize the contributions of TA/CD as an important element of budget support to directly strengthen national systems, particularly the PFM. However, the progress in PFM depends on institutional capacity and the recipient government's political will for PFM reform.

Corruption

The analysed data show no systematic negative effect of budget support on corruption and thus do not confirm this assumed risk. Budget support has a positive effect on the degree of monitoring efforts on corruption, due to the PAF. However, the evidence suggests that the prosecution and accusation of suspects is still rare, and a conviction even rarer.

Apart from that, budget support's effects on corruption are scarcely researched in the sources. Thus, it cannot be determined whether budget support reduces corruption or has no effect.

Domestic accountability and democratic governance

Budget support contributes to improvements on the supply side of domestic accountability.^{iv}

Budget support strengthens the budget process, as it is channelled through the national budget system. It substantially augments the role of the supreme audit institutions (SAIs), where improvements are found regarding the quality and quantity of audit reports. With regard to the demand side of accountability,^v long-term effects of budget support on the role of parliament and civil society could not be identified, as the described effects are weak and inconsistent.

However, due to insufficient coverage by evidence, it cannot be answered if the positive influence on the supply side of domestic accountability stems from financial or non-financial budget support inputs.

Economic performance and income poverty

The evidence shows that budget support programmes reinforce pre-existing macroeconomic stability. Repayment of domestic debt as one trigger of economic growth is also plausibly attributed to budget support programmes. The evidence describes further positive effects on economic performance, but these are not attributed to budget support alone.

Effects of budget support on income poverty cannot be presented with certainty, due to challenges in the attribution of observed effects to budget support on outcome and impact level. Concerning specific budget support inputs, only the effect of funding on income poverty is sufficiently covered. The effect is slightly positive, although no distinct effect is reported in the evidence. For all other inputs, no generalizable conclusions on effects on economic performance and income poverty can be drawn.

Service delivery and non-income poverty

Overall, budget support is effective in increasing access to public services. Funds from SBS and GBS are decisive in increasing service delivery, particularly in the education and health sectors. Yet, budget support does not prove to be effective in improving the quality of provided services or the administration responsible for delivering the services. The evidence describes positive effects on non-income poverty in

^{iv} Institutions on the supply side of domestic accountability include government institutions, such as the ministry of finance, and statistic offices expected to supply (budget) information to the general public. Supreme audit institutions (SAI) can be part of the demand as well as the supply side, and fulfil both functions, depending on their institutional and legal status.

^v The demand side of domestic accountability consists of actors such as the parliament, civil society, and the media, who demand information from government to hold the government to account.

the recipient countries, but the effects are not attributed to budget support with certainty.

Regarding specific budget support inputs, the effects of funding, policy dialogue and accompanying TA/CD on public service delivery are sufficiently covered. On the contrary, conclusions for the effects of conditionality on public service delivery are not possible. Effects of specific inputs on administration of service delivery and non-income poverty have also not been sufficiently covered.

Conclusion

This evaluation synthesis presents the most comprehensive analysis to date of existing evidence on the effectiveness of budget support, a highly controversial modality of development cooperation. The findings systematically substantiate existing knowledge and generate new insights for decision makers in German and international development cooperation.

The study finds convincingly broad evidence that budget support is indeed an effective modality in promoting important development outcomes, such as improvements in public financial management and budget processes and improved provision of public goods and services. In view of these findings it appears worthwhile for donors – including those who have largely withdrawn from budget support – to re-assess the modality. In doing so, however, particular attention will need to be paid to the remaining evidence gaps identified in this evaluation synthesis. Specifically, there is a lack of knowledge with regard to the effectiveness of budget support at outcome and impact level due to unresolved methodological challenges and problems of attribution. Also, while the evidence proves effects of budget support programmes on a general level, little is known on the specific contribution of individual, particularly non-financial, budget support inputs. The same is true for hitherto largely neglected cross-cutting topics, such as budget support's effects on gender equality and distributional effects of growth. Furthermore, a serious knowledge deficit exists with regard to potential risks of the modality, such as corruption. This lack of evidence is particularly surprising as the political debate on budget support revolves to a large extent around corruption risks.

Outlook

The remaining gaps imply that, even from an extensive analysis of existing evidence such as this and despite strong evidence for positive effects of budget support in general, only limited conclusions can be drawn with regard to specific effects attributable to individual budget support inputs. Going hand in hand, there are only limited lessons to be learned for the design and implementation of aid modalities with similar features as budget support, such as results-based approaches. This also makes it difficult to predict the impact on development outcomes achieved through budget support of a widespread withdrawal of bilateral donors from the instrument.

Future empirical work therefore needs to closely analyse the effects and causal mechanisms of specific budget support inputs, as well as budget support effects on important cross-cutting issues. Also, future work should analyse in more depth in how far potential risks of budget support materialise in practice and how those risks can be mitigated.

To account for this need of more empirical work on the topic, DEval conducts a complementary evaluation of the exit from budget support, building on the findings of this evaluation synthesis and investigating the sustainability of budget support effects when donors suspend or exit from budget support. In conjunction with the findings presented here, this evaluation contributes to forming a more complete picture of the effectiveness of budget support, the consequences when donors exit from the aid modality, and the lessons for the design and implementation of future budget support programmes and related aid instruments.

ZUSAMMENFASSUNG

Hintergrund

Um die Jahrtausendwende vollzog sich in der internationalen Entwicklungszusammenarbeit (EZ) ein Paradigmenwechsel, der nicht zuletzt durch eine wachsende Anzahl von Wirksamkeitsuntersuchungen ausgelöst wurde. Diese Untersuchungen legten nahe, dass die Entwicklungspolitik der letzten Jahrzehnte ihre häufig wechselnden, jedoch immer ambitionierten Ziele oftmals nicht erreicht hatte. Gleichzeitig deuteten sie darauf hin, dass die fehlende Wirksamkeit zumindest teilweise auf die Art und Weise zurückgeführt werden konnte, wie die Zusammenarbeit gestaltet wurde.

EZ-Ansätze waren bis dahin von einer starken Geberkontrolle und politischen Konditionalitäten begrenzter Glaubwürdigkeit geprägt. Statt die gewünschten entwicklungspolitischen Erfolge zu erzielen, gingen diese gebergelenkten und oft stark fragmentierten Ansätze häufig mit fehlenden systemischen Wirkungen auf die soziale, ökonomische und politische Entwicklung der Partnerländer einher, unter anderem verursacht durch nicht-intendierte Wirkungen. Denn diese – in der Regel projektbasierten – EZ-Ansätze machten, trotz ihrer Absicht landeseigene Institutionen und die Regierungen zu stärken, Parallelstrukturen notwendig, die zu hohen Transaktionskosten führten und die demokratische Kontrolle, die Eigenverantwortung der Partnerregierungen sowie die lokalen Kapazitäten schwächten.^{1 vi}

Neben erkannten Schwächen von projektbasierten Ansätzen standen vor allem die sogenannten Strukturanpassungsprogramme der Bretton-Woods-Institutionen in der Kritik. Im Rahmen dieser Programme vergaben der Internationale Währungsfonds und die Weltbank vergünstigte Kredite an verschuldete oder insolvente Staaten, die im Gegenzug einem vorgefertigten Reformplan zustimmen mussten. Im Rückblick zeigte sich, dass der Versuch, mittels Strukturanpassungsprogrammen die Wirtschaftspolitik in den Partnerländern zu reformieren, nicht die gewünschten Ergebnisse gebracht hatte.²

Angesichts der problematischen Erfahrungen mit konventioneller projektbasierter EZ und Strukturanpassungskrediten verstärkten Geber und Partnerregierungen ihre Anstrengungen, die Zusammenarbeit neu zu definieren und damit die

Effektivität der EZ voranzutreiben. Auf mehreren Gipfeltreffen formulierte die internationale Gemeinschaft schrittweise eine neue Agenda für eine wirksamere EZ, die 2005 in der Erklärung von Paris über die Wirksamkeit der EZ in fünf zentrale Prinzipien gefasst wurde:³

- Eigenverantwortung der Partnerregierungen
- Ausrichtung der Programme an den Strategien und Verfahren der Partnerregierungen
- Harmonisierung von Programmen und Verfahren der Geber
- Ergebnisorientierung
- Gegenseitige Rechenschaftspflicht

Zur Umsetzung dieser Prinzipien in der Praxis mussten neue Modalitäten eingesetzt werden, die weithin unter dem Begriff „Programmbasierte Ansätze“ gefasst werden. Sie erfordern die Eigenverantwortung der Partnerregierung, einen formalisierten Prozess für Geberkoordinierung, Harmonisierung der Geber sowie die verstärkte Nutzung der landeseigenen Systeme.

Den programmbasierten Ansätzen zugehörend gilt die Budgethilfe, die in der zweiten Hälfte der 2000er Jahre an Bedeutung gewann, als eine der konsequentesten Modalitäten für die Umsetzung der in der Erklärung von Paris formulierten Prinzipien. Meist als Gemeinschaftsvorhaben mehrerer Geber konzipiert, besteht das Hauptziel der Budgethilfe darin, die Implementierung der nationalen Armutsbekämpfungsstrategie der jeweiligen Partnerregierung finanziell zu unterstützen (Finanzierungsfunktion) und durch in nationalen Institutionen und Foren ausgehandelte Reformprozesse gute Regierungsführung zu fördern (Governancefunktion). Während die allgemeine Budgethilfe einen nicht zweckgebundenen Beitrag zu den Staatsausgaben der Partnerregierung darstellt, werden die Beiträge der Sektorbudgethilfe an die Nutzung für einen spezifischen Sektor geknüpft.

Zu diesem Zweck besteht die Budgethilfe sowohl aus finanziellen Inputs der Geber, die direkt in den Haushalt der Partnerregierung fließen wie auch aus nicht-finanziellen Inputs, die den Reformprozess unterstützen sollen. Die nicht-finanziellen Beiträge setzen sich aus drei Inputs zusammen: Konditionalitäten, Politikdialog und begleitendem Kapazitätsaufbau. Die Konditionalitäten bestehen dabei in der Regel auf zwei

^{vi} In der gesamten Evaluationssynthese werden die verwendeten Quellen in Endnoten (arabisches Zahlensystem) und Erklärungen sowie Beispiele in Fußnoten (römisches Zahlensystem) aufgeführt.

Ebenen: den sogenannten Underlying Principles, die Grundlage für die Vergabe allgemeiner Budgethilfe darstellen; und sogenannte Performance Assessment Frameworks, in deren Rahmen zwischen Partnerregierungen und den Gebern Zielwerte und Indikatoren verhandelt werden, die dem Monitoring der Umsetzung der nationalen Entwicklungsstrategie und der verschiedenen Reformprozesse dienen. Der Politikdialog zwischen Partnerregierungen und Gebern findet sowohl auf technischer Ebene (insbesondere wird hier die Erfüllung der im Performance Assessment Framework festgelegten Ziele bewertet), als auch auf einer übergreifenden, strategischen Ebene statt. Der begleitende Kapazitätsaufbau, als dritter nicht-finanzieller Input, dient der Stärkung von Kapazitäten der Regierung und Institutionen im Partnerland und soll damit die Mittelverwendung der Partnerregierung positiv beeinflussen. In der Kombination von finanziellen und nicht-finanziellen Inputs sollen Wechselwirkungen entstehen, die die entwicklungspolitische Effektivität beider Inputs erhöhen.

Um die Budgethilfe entspann sich bald eine kontroverse Debatte in der entwicklungspolitischen Fachöffentlichkeit. In dieser Debatte dominieren zwei Ansichten: Unter den Befürworterinnen und Befürwortern gilt speziell die allgemeine Budgethilfe als besonders geeignetes Instrument zur Intensivierung von Eigenverantwortung der Partnerregierungen, Ausrichtung an Partnerstrukturen und Harmonisierung unter den Gebern. Durch die Kombination aus externen Finanzmitteln und der Förderung von sektorspezifischen und übergreifenden Reformen wird von systemischen Wirkungen auf das Partnerland ausgegangen. Das Ziel der Armutsreduzierung wird dabei indirekt über die Unterstützung von Aktivitäten der Partnerregierung und die Nutzung von landeseigenen Systemen realisiert. Demgegenüber steht die Sorge um treuhänderische und politische Risiken der Budgethilfe, sollte die Qualität der Systeme auf Partnerseite unzureichend sein bzw. sich nach anfänglicher Prüfung verschlechtern. Als treuhänderische Risiken nennen Kritikerinnen und Kritiker eine Veruntreuung der mit der Budgethilfe bereitgestellten Gelder, zum Beispiel durch Korruption, wie auch die Fehlallokation von Mitteln durch die Partnerregierung, beispielsweise in nicht-armutsrelevante Sektoren. Aufseiten der politischen Risiken werden mögliche negative Effekte auf die interne Rechenschaftslegung

der Regierung gegenüber Parlament und Bevölkerung (erklärt durch eine primäre Unterstützung der Exekutive) und verminderte Anreize zur Generierung eigener Einnahmen befürchtet.

Während zu Beginn der Debatte stärker die Finanzierungsfunktion der Budgethilfe als primäres Ziel im Vordergrund stand, rückte später, insbesondere unter den bilateralen Gebern, die Governance- bzw. Reformfunktion der Modalität in den Fokus.⁴

Mit wachsender Bedeutung der Budgethilfe wurden auch verstärkt Evaluierungen durchgeführt, die die Wirksamkeit der Modalität anhand einer von dem Entwicklungsausschuss der OECD (Development Assistance Committee, DAC) anerkannten Wirkungslogik überprüfen sollten.^{vii} Diese Wirkungslogik beschreibt die angestrebten Ergebnisse von Budgethilfe, die mit Hilfe der Umsetzung der oben beschriebenen Paris-Prinzipien der Wirksamkeit von Entwicklungszusammenarbeit erreicht werden sollen. Es wird eine stärkere Ausrichtung der Geber an den Politiken und Systemen der Partnerregierung erwartet sowie eine stärkere Harmonisierung unter Gebern und eine bessere Vorhersagbarkeit der Mittel erhofft. Durch die Nutzung nationaler Systeme verspricht man sich zudem eine Senkung der Transaktionskosten. Es werden erhöhte Regierungsausgaben für soziale Sektoren, ein gestärktes öffentliches Finanzwesen, weniger Korruption, eine erhöhte innerstaatliche Rechenschaftspflicht und demokratisches Regieren als Ergebnis der Budgethilfe angenommen. Auf Ebene von Outcomes und Impact werden eine verbesserte wirtschaftliche Leistung und reduzierte Einkommensarmut sowie eine vermehrte Bereitstellung öffentlicher Güter und eine verringerte Nicht-Einkommensarmut erwartet.

Einige dieser (meist von mehreren Gebern gemeinsam durchgeführten) Evaluierungen bescheinigen der Budgethilfe positive Effekte in Bezug auf ihre Finanzierungsfunktion und attestieren ihr positive Wirkungen im Bereich des öffentlichen Finanzmanagements. Die übergeordnete kritische Diskussion über die Budgethilfe blieb von diesen Evaluierungen allerdings weitgehend unbeeinflusst, da diese als Einzelergebnisse wahrgenommen wurden, die die Perspektive eines einzelnen Gebers oder einer individuellen Organisation widerspiegeln. Um diese Wahrnehmung zu ändern, wurden die Ergebnisse

^{vii} Die Wirkungslogik ist im „Comprehensive Evaluation Framework (CEF)“ beschrieben, der als Teil des methodischen Ansatzes zur Evaluierung der Budgethilfe unter der Federführung des OECD DAC und der Europäischen Kommission entwickelt wurde.

einzelner Evaluierungen in Synthesestudien zusammengefasst. Jedoch basieren diese meist auf einer begrenzten Anzahl an Fällen und lassen nicht immer erkennen, wie flächendeckend die dargestellten Ergebnisse durch Evidenz gestützt sind.

In den letzten Jahren hat die Akzeptanz für Budgethilfe in den Geberländern deutlich abgenommen, woraufhin viele Geberregierungen die Hilfen ausgesetzt oder ganz gestoppt haben. Inzwischen konzentriert sich die Debatte um geeignete Finanzierungsformen meist auf neue Modalitäten, auch vor dem Hintergrund einer großen Finanzierungslücke zur Erreichung der Agenda 2030. Diskutiert wird ein breites Spektrum an Modalitäten, obwohl nicht alle als neu bezeichnet werden können. Besonders prominent werden ergebnisbasierte Modalitäten diskutiert. Die sogenannte ergebnisorientierte Hilfe (*results-based aid*) identifiziert bewertbare – und auch quantifizierbare – Leistungen und Ergebnisse. Die Auszahlung der Gelder erfolgt erst, wenn die zuvor definierten Leistungen oder Ergebnisse durch die Partnerregierung erreicht worden sind. Auf diese Weise sollen Anreize zur Zielerreichung gesetzt und befürchtete treuhänderische Risiken reduziert werden. Während manche ergebnisbasierten Modalitäten alle Auszahlungen an zuvor erbrachte Leistungen knüpfen, integrieren andere nur einzelne Komponenten. Dies ist der Fall bei Budgethilfe, die über leistungsorientierte Tranchen einzelne Auszahlungen von der Erfüllung von Indikatoren des Performance Assessment Frameworks abhängig macht. Die Ähnlichkeit in der Ergebnisorientierung, der Fokus auf Outputs und Outcomes und auch die Nutzung nationaler Systeme der Partnerregierung bei der Budgethilfe und der ergebnisorientierten Hilfe erlauben es, aus den Erfahrungen der Budgethilfe Lehren für die ergebnisbasierte Hilfe zu ziehen. Dies ist insbesondere von Relevanz, da bisher nur wenig praktische Erfahrung mit neueren Formen von ergebnisorientierter Hilfe besteht, zum Beispiel cash on delivery.⁵

Zielsetzung

Vor diesem Hintergrund widmet sich die vorliegende Evaluationssynthese der Aufgabe, die vorhandene Evidenz zur Wirksamkeit von Budgethilfe systematisch aufzuarbeiten. Mit der Absicht, wichtige Wissenslücken zu schließen, wurde eine große Anzahl an Quellen (95) zugrunde gelegt und gezielt unterschiedliche Perspektiven auf die Wirkungen der

Budgethilfe analysiert. Zudem wurde die qualitative Heterogenität der einzelnen Berichte mithilfe einer Qualitätsbewertung untersucht und konsequent die Abdeckung durch Evidenz für alle Ergebnisse dargestellt.

Zielsetzung der Evaluationssynthese ist es, neben Empfehlungen für Budgethilfe-Programme auch Implikationen für die Ausgestaltung und die Implementierung zukünftiger Finanzierungsinstrumente herauszuarbeiten, die der Budgethilfe ähnliche Attribute aufweisen. Durch die systematische Aufarbeitung der vorhandenen Evidenz soll außerdem auf eine Objektivierung der Debatte um die Wirksamkeit der Budgethilfe hingewirkt werden.

Zu diesem Zweck werden die folgenden beiden Evaluierungsfragen beantwortet:

1. Welche Effekte von Budgethilfe werden durch verlässliche Evidenz bestätigt?
2. Unter welchen Bedingungen führt Budgethilfe zu Ergebnissen?

Methode

Die vorliegende Evaluationssynthese folgt einem theoriebasierten Ansatz. Grundlage bildet hierzu die durch den OECD DAC anerkannte Wirkungslogik, die in bisherigen Evaluierungen verwendet wurde und die erwarteten Wirkungen der Budgethilfe beschreibt. Insgesamt wurden 95 Quellen ausgewertet, davon 32 Evaluierungen, 42 wissenschaftliche Veröffentlichungen und 21 Werke aus dem Bereich der „grauen“ Literatur. Die Quellen umfassen sowohl Untersuchungen der allgemeinen Budgethilfe als auch der Sektorbudgethilfe. Somit ist gewährleistet, dass eine große Anzahl an Quellen und verschiedene Perspektiven auf die Wirkungen der Budgethilfe abgedeckt und analysiert werden.

Die Ergebnisse aus überwiegend qualitativen Evaluierungen und Studien werden in Anlehnung an die hohen Standards einer systematischen Überprüfung (*systematic review*) synthetisiert. Dazu folgt die Evaluationssynthese einem dreistufigen Ansatz, der eine explizite Suchstrategie, klare Einschlusskriterien sowie eine systematische und softwaregestützte Codierung und Analyse beinhaltet. Als ein zentrales Element der

Methode wurde eine Qualitätsbewertung der Quellen durchgeführt. Die Beurteilung zeigt eine hohe Homogenität in der methodischen Qualität der Evaluationen sowie der wissenschaftlichen und „grauen“ Literatur. Die methodische Qualität der Evidenz ist ausreichend hoch, um alle Berichte als Basis dieser Evaluationssynthese zu nutzen.

Mit dem Ziel, die durch verlässliche Evidenz gesicherten Effekte der Budgethilfe herauszuarbeiten, wurde zum einen anhand der Abdeckung in den Quellen zwischen empirisch belegten und nicht hinreichend belegten Effekten unterschieden, wobei sowohl die Effekte des gesamten Budgethilfe-Programms wie auch Effekte der einzelnen (finanziellen und nicht-finanziellen) Inputs betrachtet wurden. Ergebnisse wurden nur dann als evident in die Evaluationssynthese aufgenommen, wenn sie von mindestens zehn verschiedenen Quellen beschrieben werden. Zum anderen wurden die Quellen anhand der Qualitätsbewertung in zwei Kategorien – „best evidence“ und „second best evidence“ – eingeteilt. Diese Differenzierung wurde für einen Robustheitstest verwendet, der positiv ausfällt, sofern Ergebnisse durch Quellen aus der Kategorie „best evidence“ oder aus beiden Kategorien gestützt sind.

Ergebnisse

Die hier zusammengefassten Ergebnisse basieren auf der qualitativen und quantitativen Analyse der Budgethilfe-Wirkungen entlang der erwarteten Effekte aus der durch den OECD DAC anerkannten Wirkungslogik.

In dieser Zusammenfassung der Ergebnisse wird unterschieden zwischen Ergebnissen, die durch häufige und konsistente Nennung in den Quellen als verlässlich angesehen werden können, und solchen, die nicht durch ausreichende Evidenz untermauert sind, da sie nur selten angeführt werden. Eine Übersicht aller empirisch belegten Ergebnisse befindet sich in Tabelle 9.

Aus der Evaluationssynthese lassen sich die folgenden Hauptergebnisse ableiten:

- Die meisten erwarteten Effekte der Budgethilfe sind ausreichend durch Evidenz abgedeckt.

- Die meisten belegten Wirkungen der Budgethilfe sind positiv oder zumindest bei Präsenz entsprechender Kontextbedingungen positiv (zum Beispiel bei einer hohen Anzahl und Qualität gemeinsamer Geberinteressen und Erfüllung von Konditionalitäten durch die Partnerregierung).
- Die beobachteten Effekte werden meist der Wirkung eines gesamten Budgethilfe-Programms zugeschrieben. Nur selten erfolgt die Zuordnung zu einzelnen Budgethilfe-Inputs wie Finanzierung, Konditionalitäten, Politikdialog oder Begleitenden Maßnahmen.
- Eine Vielzahl der Quellen vermittelt glaubhaft die Zuordnung beobachteter Entwicklungen zur Budgethilfe auf Ebene der Outputs und Induced Outputs. Auf Ebene von Outcomes und Impacts der Wirkungslogik wird die Zuschreibung zu Budgethilfe-Programmen meist nicht mehr mit ausreichender Evidenz abgedeckt.
- Wichtige und viel diskutierte Risiken der Budgethilfe wurden von den zugrundeliegenden Quellen nur unzureichend untersucht, so dass hierzu keine Aussagen möglich sind.
- Die Stärke der Budgethilfe-Effekte ist stark abhängig von Kontextfaktoren wie institutionellen Kapazitäten und Reformwillen der Partnerregierung.

Prinzipien für die Wirksamkeit der Entwicklungszusammenarbeit^{viii}

Im Ergebnis zeigen Budgethilfe-Programme einen positiven Effekt auf die stärkere Ausrichtung der Programme an den Strategien und Verfahren der Partnerregierungen, die konsequentere Harmonisierung von Programmen und Verfahren der Geber sowie die Stärkung der Eigenverantwortung der Entwicklungsländer. Allerdings ist der Wirkungsgrad abhängig vom Umfang und von der Qualität der gemeinsamen Interessen, die einerseits zwischen den Gebern und andererseits zwischen den Gebern und Partnerregierungen bestehen. Entscheidend ist zudem die Einhaltung der mit dem Programm verbundenen Konditionalitäten durch die Partnerregierung.

Die Budgethilfe-Programme zeigen einen insgesamt reduzierenden Effekt auf die Höhe der Transaktionskosten, nicht jedoch auf die Verlässlichkeit der Zahlungen. Der Grund dafür ist zum einen die häufig verspätete Auszahlung der Finanzhilfen im Fiskaljahr, zum anderen das unkoordinierte Aussetzen von Auszahlungen infolge des Nichteinhaltens von abgestimmten

^{viii} Diese Prinzipien wurden aus der Interventionslogik der Budgethilfe abgeleitet und sind nicht identisch mit den Prinzipien der Erklärung von Paris, sondern umfassen die Ausrichtung der Programme an Strategien und Verfahren der Partner, die Harmonisierung von Programmen und Verfahren der Geber, die Eigenverantwortung der Partnerländer, verlässliche Zahlungen und eine Reduktion der Transaktionskosten.

Konditionalitäten. Im Zeitverlauf zeigt sich, dass die Transaktionskosten nach der Einführung von Budgethilfe zunächst stiegen, nach der Etablierung gemeinsamer Prozesse aber wieder sanken. Wenn sich Geber aus der Budgethilfe zurückziehen, lässt das Aussetzen von Tranchen die Transaktionskosten jedoch in der Regel wieder steigen, da nur noch geringe Summen im Verhältnis zum operativen Aufwand ausgezahlt werden.

Bei den nicht-finanziellen Inputs der Budgethilfe, insbesondere beim Politikdialog und der Konditionalität, zeigt sich eine stärkere Harmonisierung der Programme und Verfahren der Geber. Sie tragen jedoch nicht dazu bei, die Eigenverantwortung der Entwicklungsländer zu erhöhen, und die Konditionalität hat eventuell sogar einen negativen Einfluss auf die Verlässlichkeit von Zahlungen.

Zu den Wirkungen weiterer spezifischer Budgethilfe-Inputs auf die Prinzipien wirksamer EZ können keine generalisierbaren Aussagen getroffen werden, da diese nicht hinreichend durch die vorhandene Evidenz abgedeckt sind. Zum Beispiel kann nicht beantwortet werden, ob sich spezifische Budgethilfe-Inputs auf eine stärkere Ausrichtung der Geberprogramme an Verfahren der Partnerregierungen auswirken und ob nicht-finanzielle Inputs der Budgethilfe zu einer Minderung der Transaktionskosten führen. Es können außerdem keine Schlussfolgerungen im Hinblick auf die Wirkung von technischer Zusammenarbeit/Entwicklung von Kapazitäten auf alle fünf Prinzipien gezogen werden.

Regierungsausgaben

Budgethilfe, dabei vor allem die Finanzhilfe, führt zu einem Ansteigen öffentlicher Ausgaben, besonders deutlich in den sozialen Sektoren Bildung und Gesundheit. Ein debattiertes Risiko, nämlich dass Budgethilfe den Anreiz für die Erhöhung von Inlandseinnahmen reduziere und diese damit verdränge (der sogenannte Crowding-out-Effekt), wird durch die Evaluationssynthese nicht bestätigt. Tatsächlich lässt sich nicht beobachten, dass Budgethilfe die Mobilisierung von Inlandseinnahmen reduziert. Gleichzeitig zeigt sich kein oder nur ein sehr geringer Anstieg der Inlandseinnahmen durch Reformen des öffentlichen Finanzwesens, was die Nachhaltigkeit der Effekte von Budgethilfe auf die (armutsrelevanten) öffentlichen Ausgaben beeinträchtigt.

Während die Wirkungen von finanziellen Budgethilfe-Inputs in den der Evaluationssynthese zugrundeliegenden Quellen häufig genannt werden, finden sich kaum Beschreibungen über die Wirkungen von nicht-finanziellen Inputs auf die Regierungsausgaben. Aufgrund dieser defizitären Abdeckung in der Evidenz konnte nicht untersucht werden, ob sich nicht-finanzielle Inputs entsprechend der Wirkungslogik positiv auf armutsorientierte Regierungsausgaben auswirken. Deshalb kann auch für zukünftige Programme oder verwandte Instrumente wenig gelernt werden.

Öffentliches Finanzwesen

Die Evaluationssynthese belegt, dass sich Budgethilfe-Programme positiv auf die Qualität des öffentlichen Finanzwesens der Partnerregierungen auswirken, insbesondere auf die Budgetplanung und -formulierung sowie die Verständlichkeit und Transparenz des Budgets. Diese Wirkung erfährt in den zugrunde gelegten Quellen die größte Abdeckung. Die Verbesserungen sind hauptsächlich zurückzuführen auf die nicht-finanziellen Inputs Politikdialog und die Kapazitätsentwicklung, da in beiden Bereichen ein Fokus auf die Verbesserung des öffentlichen Finanzwesens gelegt wird. Speziell spätere Evaluationen erkennen an, dass im Rahmen der Budgethilfe durchgeführte Maßnahmen nationale Systeme, vor allem das öffentliche Finanzwesen, stärken. Das Ausmaß des Fortschritts im öffentlichen Finanzwesen hängt allerdings von den institutionellen Kapazitäten und dem politischen Reformwillen aufseiten der Partnerregierung ab.

Korruption

Die Synthese zeigt keinen systematischen negativen Effekt von Budgethilfe auf Korruption und bestätigt damit ein solches befürchtetes Risiko nicht. Budgethilfe wirkt sich vielmehr aufgrund der Durchführung einer jährlichen Leistungskontrolle (Performance Assessment Framework) positiv auf die Monitoring-Anstrengungen von Korruption aus. Allerdings deuten die genutzten Quellen gleichzeitig daraufhin, dass eine Strafverfolgung und Anklage von Verdächtigen weiterhin selten bleibt und eine Verurteilung noch seltener stattfindet.

Von diesen Effekten abgesehen, wird der Effekt von Budgethilfe auf Korruption in den Studien kaum untersucht. Es kann damit nicht beantwortet werden, ob Budgethilfe Korruption reduziert oder kein Effekt besteht.

Innerstaatliche Rechenschaftspflicht und demokratisches Regieren

Budgethilfe stärkt die „Angebotsseite“^{ix} der innerstaatlichen Rechenschaftspflicht, insbesondere durch die Stärkung des Budgetprozesses über die Nutzung der nationalen Systeme. Die Rolle der obersten Rechnungskontrollbehörden wird durch Budgethilfe erheblich gestärkt. Dies zeigt sich hauptsächlich in der Qualität und der Quantität von Prüfungsberichten. Im Hinblick auf die Nachfrageseite^x der innerstaatlichen Rechenschaftspflicht konnten Langzeiteffekte der Budgethilfe auf die Rolle des Parlaments und der Zivilgesellschaft nicht identifiziert werden, da in den Quellen nur schwache, inkonsistente Effekte beschrieben werden.

Aus welchen Budgethilfe-Inputs (finanzieller oder nicht-finanzieller Natur) die positiven Effekte auf die Angebotsseite der innerstaatlichen Rechenschaftspflicht resultieren, lässt sich aufgrund geringer Abdeckung in den Quellen nicht hinreichend sicher herleiten.

Wirtschaftsleistung und Einkommensarmut

Budgethilfe verstärkt bestehende makroökonomische Stabilität. Die Tilgung von Inlandsschulden wird als ein Katalysator des Wirtschaftswachstums außerdem den Effekten der Budgethilfe zugesprochen. Weitere positive Entwicklungen in der Wirtschaftsleistung werden ihr nicht eindeutig zugewiesen.

Effekte der Budgethilfe auf Einkommensarmut können ebenfalls nicht mit ausreichender Sicherheit dargestellt werden. Dies ist darauf zurückzuführen, dass sich auf dieser Ebene die Zuschreibung beobachteter Effekte auf die Budgethilfe schwierig gestaltet. Mit Blick auf die Wirksamkeit spezifischer Budgethilfe-Inputs liegt eine hinreichende Evidenz lediglich für die Wirkung finanzieller Inputs auf die Einkommensarmut vor; es lässt sich ein leicht positiver, wenn auch kein eindeutiger Trend konstatieren. Für alle weiteren Inputs können aufgrund

zu geringer Evidenz keine generalisierbaren Schlussfolgerungen zu den Auswirkungen auf die Wirtschaftsleistung und die Einkommensarmut getroffen werden.

Bereitstellung öffentlicher Leistungen und Nicht-Einkommensarmut

Budgethilfe erhöht den Zugang zu öffentlichen Dienstleistungen. Gelder aus der allgemeinen sowie der Sektorbudgethilfe sind maßgeblich für die Ausweitung der Bereitstellung von Leistungen, vor allem in den Sektoren Bildung und Gesundheit. Allerdings konnte nicht festgestellt werden, dass Budgethilfe die Qualität der zur Verfügung gestellten Dienstleistungen oder die Verwaltung derselben verbessert. Die sich in den Empfängerländern abzeichnenden positiven Trends in der Nicht-Einkommensarmut lassen sich auf Basis der bestehenden Evidenz nicht mit Sicherheit auf die Budgethilfe zurückführen.

Hinreichende Evidenz zu Budgethilfe-Effekten auf die Erbringung von öffentlichen Leistungen liegt für die Wirkung von finanziellen Inputs, des Politikdialogs und der begleitenden technischen Zusammenarbeit/Entwicklung von Kapazitäten vor. Für die Wirkung von Konditionalität auf die Erbringung von öffentlichen Leistungen können hingegen keine generalisierbaren Schlussfolgerungen gezogen werden. Effekte von spezifischen Budgethilfe-Inputs auf die Verwaltung öffentlicher Leistungen und die Nicht-Einkommensarmut sind ebenfalls nicht durch ausreichende Evidenz abgedeckt.

Fazit

Diese Evaluationssynthese stellt den bislang umfassendsten Blick auf die existierende Evidenz zur Wirksamkeit von Budgethilfe, einer umstrittenen Modalität der Entwicklungszusammenarbeit, dar. Die Ergebnisse sichern bereits bestehendes Wissen systematisch ab und schaffen neues Orientierungswissen für Entscheidungsträger der deutschen und internationalen Entwicklungszusammenarbeit.

Zusammenfassend belegt die breite Evidenz relevante Wirkungen der Budgethilfe bei der Förderung wichtiger Entwicklungsziele, zum Beispiel bei der Verbesserung im Bereich des

^{ix} Institutionen der Angebotsseite innerstaatlicher Rechenschaftspflicht umfassen Institutionen wie das Finanzministerium und statistische Ämter, die der Öffentlichkeit Informationen zur Verfügung stellen, zum Beispiel zu Fragen des Haushalts. Rechnungskontrollbehörden können in Abhängigkeit ihres institutionellen wie rechtlichen Status der Angebots- oder Nachfrageseite angehören.

^x Die Nachfrageseite innerstaatlicher Rechenschaftspflicht umfasst Akteure wie das Parlament, die Zivilgesellschaft und die Medien, die von der Regierung Informationen zur Ablegung von Rechenschaft fordern.

öffentlichen Finanzwesens und von Haushaltsprozessen sowie bei der Bereitstellung öffentlicher Güter und Leistungen. Vor diesem Hintergrund erscheint eine erneute Prüfung der Budgethilfe als einer Finanzierungsform auch für diejenigen Geber lohnenswert, die sich weitgehend aus der Modalität zurückgezogen haben. Dabei sollte allerdings ein besonderes Augenmerk auf die in dieser Evaluationssynthese ebenfalls identifizierten verbleibenden Wissenslücken gelegt werden. Denn nach wie vor bestehen erhebliche Erkenntnisdefizite mit Blick auf wichtige Wirkmechanismen der Budgethilfe. Diese bestehen insbesondere in Bezug auf die Wirksamkeit auf der Ebene von Outcomes und Impacts aufgrund ungelöster methodischer Schwierigkeiten bei der kausalen Zuordnung von beobachteten Veränderungen. Zwar belegt die Evidenz Wirkungen von Budgethilfe-Programmen generell, liefert jedoch nur geringe Erkenntnisse zu den individuellen Wirkmechanismen der einzelnen, insbesondere der nicht-finanziellen, Budgethilfe-Inputs. Gleiches gilt für bisher vernachlässigte Querschnittsthemen wie die Gleichberechtigung der Geschlechter und Verteilungseffekte von Wachstum. Des Weiteren besteht ein gravierendes Wissensdefizit mit Blick auf potenzielle Risiken der Modalität wie Korruption. Dieser Mangel an Evidenz überrascht besonders, da sich die politische Debatte um Budgethilfe vor allem kritisch auf das Thema Korruption richtet.

Ausblick

Diese verbleibenden Wissenslücken bedeuten, dass – obwohl fraglos profunde Belege für die Wirksamkeit von Budgethilfe vorliegen – auch bei breitestmöglicher Berücksichtigung der bislang bestehenden Evidenz nur bedingt Schlussfolgerungen zu den Wirkmechanismen spezifischer Budgethilfe-Inputs gezogen werden können. Damit einhergehend können auch nur in begrenztem Umfang Lehren für die Ausgestaltung und die Implementierung von Finanzierungsmodalitäten mit ähnlichen Eigenschaften wie die Budgethilfe, zum Beispiel ergebnisbasierte Ansätze, gezogen werden. Ebenso kann auf Grundlage der vorhandenen Evidenz nicht ohne weiteres beantwortet werden, welche Folgen der in den letzten Jahren beobachtete weitgehende Ausstieg vieler bilateraler Geber aus der Budgethilfe für bereits erreichte Wirkungen hat und haben wird.

Künftige empirische Arbeiten sollten daher detailliertere Untersuchungen zu den Effekten und zugrunde liegenden

Wirkmechanismen von einzelnen Budgethilfe-Inputs wie auch zu Wirkungen in wichtigen Querschnittsbereichen anstellen. Zudem muss die weiterführende Forschung eingehender analysieren, inwiefern sich potenzielle Risiken der Budgethilfe tatsächlich realisieren und gegebenenfalls gemindert werden können.

Das DEval stellt deshalb eine komplementäre, auf den durch diese Evaluationssynthese generierten Erkenntnissen aufbauende, Evaluierung zum Ausstieg aus der Budgethilfe an. Diese Evaluierung untersucht die Nachhaltigkeit der in dieser Studie belegten Budgethilfe-Effekte nach dem Aussetzen von Zahlungen beziehungsweise dem vollständigen Ausstieg aus der Budgethilfe. In Verbindung mit den Erkenntnissen der Evaluationssynthese verfolgt sie das Ziel, zu einem Gesamtbild der Wirksamkeit von Budgethilfe, der Konsequenzen des Ausstiegs sowie der zu ziehenden Schlussfolgerungen für die wirksame Ausgestaltung und Implementierung von Budgethilfe und verwandten Instrumenten, beizutragen.

CONTENTS

Executive Summary	vii
Zusammenfassung	xiii
Abbreviations	1

1. Introduction 2

1.1	Budget support as an aid modality	3
1.2	The exit from budget support and the advent of new financing modalities	4
1.3	First findings on budget support and lack of impact on the political debate	6
1.4	The need for comprehensive synthesis work	7

2. Synthesis design and methodology 9

2.1	Evaluative challenges	10
2.2	Design and methodology	12
2.2.1	A search strategy based on explicit criteria	14
2.2.2	Relevance criteria for inclusion of sources	14
2.2.3	Systematic coding	17
2.2.4	Data analysis and reporting	19

3. Body of evidence 21

3.1	Characteristics of the sources of evidence	22
3.2	Reliability of evidence: methodological quality	23
3.3	Aggregate number of sources and codings	28
3.4	Strength of budget support effects on output and outcome level	31

4. Evidence on budget support 32

4.1	Aid-effectiveness principles	36
4.1.1	Expected effects	36
4.1.2	Alignment	37
4.1.3	Harmonization	38
4.1.4	Ownership	39
4.1.5	Predictability	40
4.1.6	Transaction costs	41
4.2	Government expenditure	42
4.2.1	Expected effects	43
4.2.2	Level of public spending	43
4.2.3	Composition of public spending	44
4.2.4	Domestic revenue mobilization	45
4.3	Public financial management	47
4.3.1	Expected effects	47
4.3.2	Budget support programmes in general	48
4.3.3	Financial and non-financial inputs	50
4.4	Corruption	51
4.4.1	Expected effects	51
4.4.2	Effects on corruption	53
4.5	Domestic accountability and democratic governance	53
4.5.1	Expected effects	54
4.5.2	Supply side of domestic accountability	54
4.5.3	Demand side of domestic accountability	55
4.5.4	Democratic governance	57
4.6	Economic performance and income poverty	58
4.6.1	Expected effects	58
4.6.2	Macroeconomic performance	58
4.6.3	Income poverty	60
4.7	Service delivery and non-income poverty	61
4.7.1	Expected effects	61
4.7.2	Public service delivery	61
4.7.3	Non-income poverty	63

5. Conclusion 65

5.1	Systematic approach	66
5.2	Summary of findings	66
5.2.1	Coverage of budget support effects	66
5.2.2	Established evidence	68
5.3	Interpretation and recommendations	70

6. References 72

Synthesis literature	73
Works cited	79

7. Annex 83

7.1	Three-step approach	84
7.2	Quality criteria evaluation reports (including scaling system)	85
7.3	Overview on budget support programmes	87

8. Endnotes 88

Figures

Figure 1	The Comprehensive Evaluation Framework for budget support	11
Figure 2	Comprehensive code system	18
Figure 3	Frequency of codings by analytical level and document type	28
Figure 4	Aggregate number of sources per outcome	29
Figure 5	Aggregate number of codings per outcome	30
Figure 6	Coverage of budget support effects	67
Figure 7	Expectation and reality of budget support effects	69

Tables

Table 1	Criteria for quality assessment of evaluation reports	16
Table 2	Criteria for quality assessment of academic and grey literature	17
Table 3	Characteristics of evaluations reports	22
Table 4	Characteristics of academic and grey literature	23
Table 5	Quality assessment of evaluation reports	24
Table 6	Quality assessment of academic and grey literature	26
Table 7	Aggregated effects of specific budget support inputs on all outputs, outcomes and impacts	31
Table 8	Coverage of evidence for all budget support effects	34
Table 9	Evidence on the effectiveness of budget support	35
Table 10	Codes on aid-effectiveness principles	36
Table 11	Budget support effects on alignment	37
Table 12	Budget support effects on harmonization	38
Table 13	Budget support effects on ownership	39
Table 14	Budget support effects on predictability	40
Table 15	Budget support effects on transaction costs	41
Table 16	Government expenditure codes	42
Table 17	Budget support effects on level of public spending	43
Table 18	Budget support effects on composition of public spending	44

Table 19	Increased spending into pro-poor sectors	45
Table 20	Budget support effects on domestic revenue mobilization	46
Table 21	Public financial management codes	47
Table 22	Budget support effects on PFM	48
Table 23	Specific improvements in PFM	49
Table 24	Corruption codes	51
Table 25	Budget support effects on corruption	52
Table 26	Domestic accountability and democratic governance codes	54
Table 27	Budget support effects on the supply side of domestic accountability	55
Table 28	Budget support effects on the demand side of domestic accountability	56
Table 29	Budget support effects on governance	57
Table 30	Economic performance and income poverty codes	58
Table 31	Budget support effects on macroeconomic performance	59
Table 32	Budget support effects on income poverty	60
Table 33	Service delivery and non-income poverty codes	61
Table 34	Budget support effects on service delivery	62
Table 35	Budget support effects on non-income poverty	63

Boxes

Box 1	The debate on budget support	5
Box 2	New finance modalities	6
Box 3	No evidence of “crowding out” of domestic revenue by budget support	44
Box 4	Evidence on accompanying TA/CD	50

ABBREVIATIONS

ADB
Asian Development Bank

AfDB
African Development Bank

CEF
Comprehensive Evaluation Framework

CSO
civil society organization

DEval
German Institute for Development Evaluation

DFID
UK Department for International Development

DRM
domestic revenue mobilization

EC
European Commission

EU
European Union

GBS
general budget support

IDB
Inter-American Development Bank

IDD
International Development Department, University of Birmingham

IEG
Independent Evaluation Group, World Bank

IMF
International Monetary Fund

LIC
low-income country

MIC
middle-income country

OECD
Organisation for Economic Co-operation and Development

OECD DAC
OECD Development Assistance Committee

PAF
Performance Assessment Framework

PBA
programme-based approach

PFM
public financial management

QC
quality criterion

SAI
supreme audit institution

SAP
structural adjustment programme

SBS
sector budget support

TA/CD
technical assistance/ capacity development



1.

INTRODUCTION

1.1 Budget support as an aid modality

Around the turn of the millennium, international development cooperation underwent a paradigm shift. Mounting evidence showed that over the past decades development aid had failed to achieve its frequently shifting, yet always ambitious, goals. The evidence also suggested that the ineffectiveness had at least partly to do with the way aid was provided. Against this backdrop, Western donors began to embark on the Paris Agenda on Aid Effectiveness.^{xi}

Past approaches to development cooperation were based on projects tightly controlled by donors and on political conditionality often of limited credibility.^{xii} In many cases they failed to produce the intended development outcomes. Instead, donor-driven and often highly fragmented aid projects had in many cases proved detrimental to social and economic development in aid-recipient countries. This form of development cooperation aimed to strengthen domestic institutions and government performance, but instead established additional layers of aid-management institutions, that increased transaction costs, and undermined democratic control, ownership and local capacities.¹²

Besides shortcomings in conventional development projects, such as parallel systems outside the government's budget, and low disbursement rates, criticism has also been directed to the so-called structural adjustment programmes (SAP) provided by the International Monetary Fund (IMF) and the World Bank. They were widely used during the 1980s and early 1990s. They comprise concessional loans to indebted or bankrupt countries linked to the recipient government's commitment to implement a pre-defined set of economic and political conditionalities based on the so-called Washington Consensus. In retrospect, these SAPs did not achieve their desired objective to reshape economic governance in the recipient countries. Critics have particularly stressed the intrusive nature of the programmes' conditionality approach, which was assessed as ineffective and undermined more partnership-based aid relations.¹³

Based on the problematic experience with conventional

project aid and policy-based lending in the 1980s and 1990s, donors and recipients were committed to rearranging aid relations and increasing the effectiveness of development cooperation. The international community formulated a set of principles for more effective aid and, in 2005, signed the Paris Declaration on Aid Effectiveness, which enshrined the following principles:¹⁴

- ownership and responsibility by recipient governments
- alignment of donor support with recipient priorities and systems
- harmonization of procedures and approaches among donors
- results-orientation
- mutual accountability

To implement these principles in practice, new aid modalities had to be adopted, commonly subsumed under the term programme-based approaches (PBAs). These PBAs are defined by the Development Assistance Committee of the Organisation for Economic Co-operation and Development (OECD DAC) as “a way of engaging in development co-operation based on the principles of co-ordinated support for a locally owned programme of development, such as a national development strategy, a sector programme, a thematic programme or a programme of a specific organisation”¹⁵. PBAs are characterized by the following features:

- (i) leadership by the host country or organization
- (ii) a single comprehensive programme and budget framework
- (iii) a formalized process for donor co-ordination and harmonization of donor procedures for reporting, budgeting, financial management and procurement
- (iv) efforts to increase the use of local systems for programme design and implementation, financial management, and monitoring and evaluation.

Budget support (BS) was one of the most significant forms of PBA. Usually provided jointly by multiple donors in support of recipient governments' national poverty reduction and development strategy, this new aid modality soon became

^{xi} In the entire evaluation synthesis, sources are listed in endnotes (Arabic numeral system) while explanations or illustrations are given in footnotes (Latin numeral system).

^{xii} For a recent review of political conditionality and its diversification beyond aid, see Koch (2015).

increasingly popular among both bilateral and multilateral donors.¹⁶ The OECD defines budget support as “a method of financing a partner country’s budget through a transfer of resources from an external financing agency to the partner government’s national treasury.”¹⁷ While general budget support (GBS) represents a non-earmarked contribution to support government spending, in sector budget support (SBS) the funds are usually earmarked for utilisation in a specific sector. In addition to transferring financial resources to the recipient government’s budget,¹⁸ the standard package of budget support comprises non-financial elements of policy dialogue, conditionality, and technical assistance and capacity development (TA/CD) to support the reform process of the partner country. Budget support conditionality consists of two pillars: underlying principles and a Performance Assessment Framework (PAF). Underlying principles form the conditionality basis for providing general budget support and include principles such as macroeconomic stability, commitment of the government to implement national development plans and reforms in PFM, and adherence to democratic principles and human rights. PAFs are negotiated among donors and partners and consist of performance indicators that measure the government performance in specific areas of the reform process. Disbursements of budget support instalments are linked to performance in the PAF.¹⁹ The policy dialogue between donors and partners takes place on technical level, in particular to assess the partner government’s performance in PAF indicators, and on a strategic, higher-ranking level. TA/CD as third non-financial input, serves to strengthen the capacities of the partner government and institutions to increase the effectiveness and efficiency in the use of the financial input. The combination of financial and non-financial inputs is expected to have reciprocal effects, increasing the effectiveness of both financial and non-financial inputs in their impact to reduce poverty.

Drawing on lessons from previous policy-based lending modalities,^{xiii} the World Bank, the European Union (EU) and some larger member states set up the first multi-donor budget support programmes around the turn of the millennium. Budget support gained importance towards the second half of the 2000s. Donors expressed high expectations with regard to the new aid instrument and disbursed increasing shares of aid in the form of budget support.²⁰ Together with the European Commission (EC), bilateral donors (e.g. the UK and the Netherlands) were frontrunners in the early phase of budget support. They were followed by other member states (among them Sweden, Denmark, Finland, and Germany), which joined the budget support programmes but took a more selective approach based on stricter governance indicators. With budget support in full swing, EU member states and the EC spent roughly EUR 15.5 billion as general budget support (GBS) between 2006 and 2010.^{xiv} Despite ambitious targets to channel up to 50% of government-to-government assistance through country systems, and making increased use of budget support and other PBAs,²¹ neither the EC nor bilateral champions of the instrument ever achieved such shares.^{xv}

1.2 The exit from budget support and the advent of new financing modalities

Towards the end of the 2000s, enthusiasm for the new aid instrument faded, particularly among bilateral donors, and the debate on budget support became increasingly politicized.²²

^{xiii} Most importantly, donors reformed their conditionality approach. “Instead of imposing conditionality and demanding policy alignment from recipient countries, donors were now expected to draw their conditionality from a comprehensive, nationally elaborated strategy on development and poverty education” (Knoll 2008: 2).

^{xiv} Data obtained from the OECD DAC CRS databank.

^{xv} For additional information on the size of budget support programmes in selected countries, see annex 7.3.

Box 1 The debate on budget support

Budget support, in particular GBS, is seen as a particularly suitable instrument to intensify ownership of the partner governments, alignment with national structures of the partner and harmonization among donors. In the combination of financial inputs (financing function) and the support of sector-specific and cross-sectoral reforms (governance function), systemic impacts on the partner country are expected. Through a harmonized contribution of several donors, geared towards the priorities of the partner country, a higher predictability of aid is expected. The reduction of poverty is to be achieved indirectly in supporting the activities of the partner government and the use of national systems of the partner.

In contrast, the debate also revolves around fiduciary and political risks of budget support in case the quality of the partner countries' systems is insufficient or changes to the negative. As fiduciary risks, the debate discusses the misappropriation of budget support funds, such as corruption, as well as a misallocation of funds in non-poverty relevant sectors. On the side of political risks, budget support opponents fear negative effects on domestic accountability due to the one-sided support of the executive and less incentives to mobilize domestic revenues.

Also, there was a controversy over the ultimate goals of the instrument. In particular during early stages of budget support programmes, it was argued that the prime – if not sole – objective of budget support should be to finance recipient countries' poverty-reduction strategies. Later, mostly among bilateral donors, budget support's function to promote good governance through its conditionality, policy dialogue, and accompanying measures shifted to the centre of interest.²³ This change in the aim hierarchy led, in combination with more sceptical risk assessments, to a more critical stance towards budget support.²⁴

Together with feared risks, political difficulties in various recipient countries, in particular corruption cases and human rights violations, led to partly uncoordinated suspensions of budget support payments by individual donors.²⁵ Arguably, with the change towards more conservative governments, European donors became increasingly concerned about fiduciary risks and started attaching stricter political conditionality to budget support.²⁶ The financial volume provided as GBS by European member states and the EC fell to EUR 4.2 billion between 2011 and 2013.^{xvi} Today, many bilateral donors, such as the Netherlands and the UK, have suspended or stopped GBS. On the request of its member states, the EC has imposed stricter entry criteria for GBS in its new budget support guidelines.²⁷ With only moderate decreases in the overall disbursement figures for all forms of budget support,^{xvii}

the EC has shifted from GBS to SBS and 82% of ongoing EU budget support programmes are now being provided as Sector Development Contracts. GBS (i.e. Good Governance and Development Contracts as well as State Building Contracts directed towards fragile states) plays a minor role, with 4% and 6% respectively.²⁸

In bilateral donors' quest to reduce fiduciary risk, new finance modalities tying aid disbursements even closer to programme performance and results have gained importance in recent years.²⁹ Such results-oriented modalities, however, have a close linkage to performance-based tranches of budget support disbursements, aiming to incentivize partner governments to achieve the desired outputs or outcomes by paying them upon delivery rather than up front.³⁰

^{xvi} Data obtained from the OECD DAC CRS databank.

^{xvii} In 2014, the EC provided EUR 1.6 billion (21% of total EuropeAid disbursements) as budget support (see EC, 2015).

Box 2 New finance modalities

With the adoption of the sustainable development goals (SDG), the discussion on effective financing for development and innovative finance modalities gained new momentum. According to the UN, the gap for financing the SDGs is estimated to be around 2.5 trillion annually.³¹

Innovative finance approaches cover a broad spectrum of modalities. One focus is on the mobilization of additional private capital, for example through government guarantees or blending. Another prominent debate evolves around results-oriented modalities. The so-called results-based aid defines measurable outputs and outcomes, and financial means are disbursed once the agreed output or outcome is achieved. While some modalities link all payments to the achievement of predefined results, others only incorporate results-oriented components. This is the case for SBS and GBS, using variable performance tranches to link single disbursement tranches to the success in defined PAF indicators. Apart from the results-orientation, both results-based aid and budget support have a strong focus on outputs and outcomes and use the partner countries' national systems.³²

This similarity of the modalities allows to derive lessons from the case of budget support for results-based aid – even more so as practical experience with newer forms of results-based aid, such as cash on delivery, is so far rare.

1.3 First findings on budget support and lack of impact on the political debate

In line with the growing use of budget support during the first decade of the millennium, efforts to assess the effects of budget support were undertaken early on.³³ As early as 2001, a group of 20 donors, among them Germany, launched an evaluability study of GBS, which developed a first methodological framework for the evaluation of budget support.³⁴ This framework subsequently guided a first major effort to evaluate the effectiveness of budget support undertaken in 2004/2005 by a consortium of donors and coordinated by the OECD DAC,³⁵ with the aim of evaluating to what extent and under which circumstances budget support had generated effects in seven country cases (Burkina Faso, Malawi, Mozambique, Nicaragua, Rwanda, Uganda, and Vietnam).

A synthesis of these country evaluations³⁶ finds positive effects of budget support on aid management (harmonization and alignment), the quality of the policy dialogue conducted between donors and recipient governments, and the management of public finances in recipient countries. The impact of budget support, e.g. on economic growth, poverty

reduction, or social sector service delivery, could not be evaluated as envisaged in the inception report of this evaluation effort, due to unresolved methodological challenges (in particular regarding the attribution of effects to budget support on the outcome and impact level).³⁷

With increasing pressure on governments, mainly from parliaments in donor countries, to demonstrate results of development cooperation at impact level, further efforts were thus made to improve on the methodology for evaluating budget support. Under the auspices of the EC, a working group consisting of the budget support donors developed an intervention logic for budget support programmes³⁸ that aimed to cover the entire causal chain, from the input to the impact level. The intervention logic describes the expected effects of budget support. The achievement of those effects is supported by implementing the principles of effective aid, such as alignment to recipient governments' policies and systems, harmonization among donors and increased predictability in the disbursement of funds. By using the systems of the recipient country, also a reduction in transaction costs is expected. Expected effects include increased government spending in social sectors, strengthened PFM, less corruption, and improvements in domestic accountability and

democratic governance. On outcome and impact level, the expected effects are improved macroeconomic performance and reduced income poverty, as well as an increase in service delivery of public goods and reduced non-income poverty.

This enhanced framework was first piloted between 2009 and 2011 in three countries (Mali, Tunisia, and Zambia^{xviii})³⁹ and has since been applied in a revised form in evaluations in various countries (also see Section 2.1).⁴⁰ In general terms, these evaluations claim clearly positive effects of budget support's financing function, and some results, however mixed, with respect to the instrument's effectiveness in fostering good governance and policy making.

To the disappointment of the instrument's proponents, however, these findings had little impact on the political debate related to budget support. This could possibly be attributed to remaining doubts regarding the methodology for evaluating budget support and thus the validity of the results. Methodological difficulties in evaluating budget support will be discussed in Section 2.1.

1.4 The need for comprehensive synthesis work

Acknowledging the limited impact of individual evaluations on the quality of the debate on budget support in donor countries, a number of efforts were undertaken to lift the debate onto a more solid evidence base through synopsis of various studies and evaluations.⁴¹ While these synthesis studies overall confirmed the evaluation findings, they were not able to change the debate on budget support, as political support for the instrument continued to dwindle in most Western donor countries. The lack of impact on the debate was arguably due to the fact that these studies were based on a limited number of studies and took specific perspectives (EC, bilateral, donor, individual aid agency). Moreover, synthesis studies tended to focus on presenting the results in a structured way without necessarily identifying which of the presented effects are widely backed by evidence.

Against this background, this evaluation addresses the task of systematically reviewing existing evidence related to the effectiveness of budget support, as well as investigating the consequences on partner countries and donors of an exit from the instrument. This report presents the findings of the conducted evaluation synthesis.^{xix} The findings are directed towards policy makers in Germany, other bilateral donors, and multilateral institutions, and implementing agencies that design and implement budget support programmes as well as related financing instruments.

This evaluation synthesis aims to provide lessons learned for the design and the implementation of future budget support and related financing modalities that, in a similar way to budget support, aim at fostering structural impact and supporting reform oriented partners. Also, the evaluation synthesis aims to contribute to a more objective discussion about the effectiveness of budget support by putting its main focus on the identification of findings that are covered by reliable evidence. To this end, the evaluation synthesis systematically reviews 95 sources, among them evaluations, academic and "grey" literature. The body of evidence is selected by using an explicit search strategy and clear inclusion/ exclusion criteria. To control for differences in the methodological quality of the different sources, all sources are rated in a quality assessment and divided into best evidence and second-best evidence. The sources are systematically coded based on the programme logic of budget support, and the codings (equals text sections) analysed to identify the effects of budget support. The findings are subject to a best-evidence robustness check, which further ensures the reliability of evidence by assessing the robustness of findings. Only those findings that are covered in 10 or more reports are finally stated in the evaluation synthesis. (For a detailed description of the methodology, see Section 2) The evaluation synthesis offers a technical analysis of the evidence and resigns intentionally from interpreting the findings to avoid bias.

^{xviii} With the participation of the Netherlands and Germany.

^{xix} The accompanying exit evaluation of budget support will be published in a separate report in mid-2017.

This evaluation synthesis reviews the existing evidence on budget support in a systematic way and aims to answer two evaluation questions:

1. Which effects of budget support are substantiated by reliable evidence?
2. Under which contextual conditions does budget support generate results?

This review of reliable budget support effects forms the basis for investigating the sustainability of budget support effects in the subsequent evaluation of the suspension of, and exit from, budget support. Together, the two products identify what is actually known about the effectiveness of budget support and the consequences that have to be faced when donors exit from the aid modality. Moreover, the products provide lessons learned about the design and implementation of future budget support programmes and related financing modalities.

This study is structured as follows. Section 2 explains the methodology and evaluation design of the evaluation synthesis. It further introduces the programme logic of budget support, the CEF being the basis for evaluating budget support and therefore also for this evaluation synthesis. The results of the quality assessment and a quantitative description of the literature analysed for this evaluation synthesis are given in Section 3. Section 4 presents the quantitative and qualitative findings of budget support effects on the different outputs, outcomes and impacts: aid-effectiveness principles, government expenditure, public financial management (PFM), corruption, domestic accountability and democratic governance, economic performance and income poverty, and service delivery and non-income poverty. Section 5 sums up the main conclusions on the effectiveness of budget support.



2.

SYNTHESIS DESIGN AND METHODOLOGY

About 15 years into implementing budget support programmes on a broad scale, knowledge about the effectiveness of the aid instrument has been gathered in numerous evaluations and academic studies. Earlier synthesis work has made important contributions to advancing this knowledge by drawing together the findings of a number of relevant studies. However, these synthesis studies did not review the substantial body of evidence in a systematic way. This section describes the systematic approach applied in the design and methodology of this evaluation synthesis.

2.1 Evaluative challenges

Budget support is a complex aid instrument in contemporary development cooperation.⁴² Defined as a financing method to provide funding to a partner country's budget by transferring resources from an external donor to the national treasury of the partner government (see Section 1),⁴³ it follows an extensive intervention logic and combines financial and non-financial inputs by multiple donors in order to achieve multiple outcomes.

There has been a long-standing methodological debate on how to adequately assess the contribution of complex budget support programmes to the envisaged changes at different levels. After a first set of country case studies conducted up until 2005 (see Section 1.3), evaluation experts under the lead

of the EC proposed a common approach to evaluating budget support. This approach includes the CEF, which describes the generic intervention logic for the aid instrument, and provides the so-called “Three-Step Approach” for evaluating budget support (see Annex 7.1).⁴⁴ This propped approach was subsequently applied in a series of pilot evaluations⁴⁵ and revised in 2012. In its latest form, the CEF (described below) served as a yardstick for the majority of multi-donor evaluations of budget support.⁴⁶

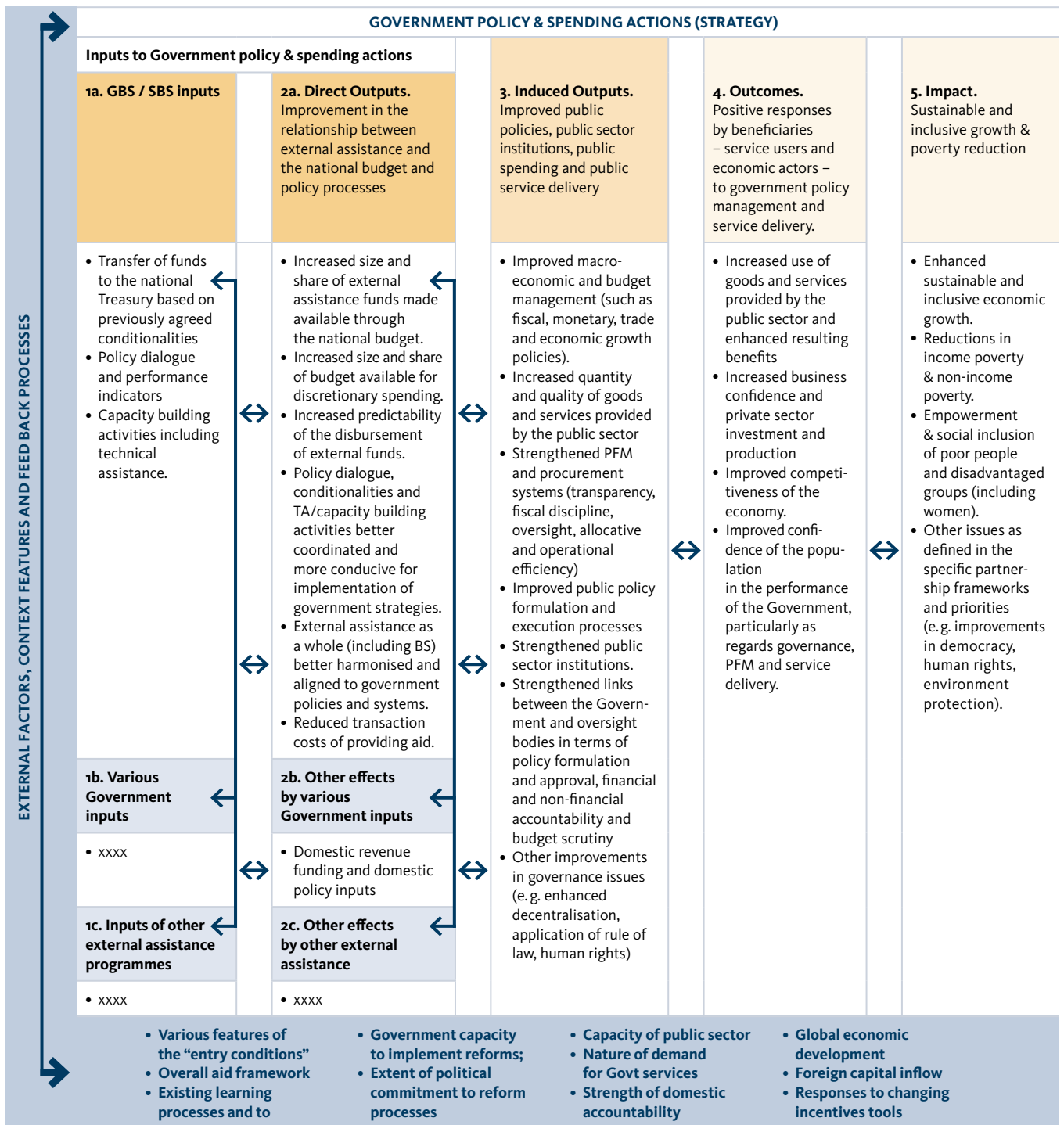
The Comprehensive Evaluation Framework (CEF)

The sequence of effects expected to result from budget support programmes has been described in an intervention logic for the instrument. Its latest version, the CEF,^{xx} illustrates the standard inputs of budget support and lists expected changes along the causal chain from direct outputs to induced outputs, outcomes, and impact (see Figure 1). If implemented in accordance with the principles of aid effectiveness, budget support is expected to produce direct outputs, e.g. increased predictability of external funds or reduced transaction costs of providing and receiving aid, and thus to increase aid effectiveness in the recipient country.⁴⁷ The CEF refers to three of the five Paris principles of aid effectiveness as direct outputs (ownership, harmonization, and alignment) and stipulates that budget support positively affects aid effectiveness through increasing predictability and reducing the transaction costs of aid in the recipient country.^{xxi}

^{xx} The CEF presented in the methodological approach to evaluating budget support by the OECD DAC (2012) illustrates the intervention logic of budget support in its most recent form. Earlier versions include the Enhanced Evaluation Framework – EEF (see EC (2008), *Methodology for Evaluations of Budget Support Operations at Country Level – Issue paper*) and the Evaluation Framework – EF (see Lawson, A./ Booth, D. (2004).

^{xxi} When referring to aid-effectiveness principles, this synthesis refers to these five principles – the three Paris principles plus the expected direct outputs of increased predictability and reduced transaction costs.

Figure 1: The Comprehensive Evaluation Framework for budget support



On the next level in the intervention logic, the CEF describes induced outputs that “... are not directly produced by budget support inputs and direct outputs, but require another actor (in this case the government) to produce them.”⁴⁸ These induced outputs are: improved macroeconomic management, increased quantity and quality of public services, and strengthened systems of PFM. This implies that changes at higher levels (from induced outputs, outcomes, and impact) are achieved indirectly by inducing changes within the recipient country’s system. Outcomes such as the increased use of publicly provided goods and services, or expected impacts such as a reduction of income and non-income poverty, depend on the recipient country’s administrative capacities and its government’s commitment to achieve the desired changes.

Budget support programmes interact in two ways with the broader context in which they are implemented. First, budget support is only one out of many drivers of change acting in parallel with other government programmes and aid interventions pursuing the same objectives. Second, changes in the wider economic, political and institutional context may influence the potential for budget support to be effective, given that the success of budget support depends on domestic context conditions, such as the quality of national policies or the efficiency of PFM systems. The programme logic formulated in the CEF assumes two major types of “driving force” to induce expected changes: “flow-of-funds effects”, which result from the financial input of budget support, and “policy and institutional effects”, caused by the non-financial inputs of policy dialogue, conditionality, and technical assistance/ capacity development (TA/CD) provided complementary to budget support funds in order to strengthen national country systems and policies.

Methodological debate

Despite the considerable resources invested in developing the methodological approach, fundamental difficulties in evaluating

budget support remain. Among others, Dijkstra and de Kemp (2015) identify as one of the main unsolved issues the challenge of defining the “correct” counterfactual for budget support, i.e. a hypothetical situation (project aid, other forms of PBAs, or a situation without aid) against which the effectiveness of budget support as an aid modality could be assessed.^{xxii} Moreover, the Three-Step Approach has been described as a “broad but weak” evaluation approach,⁴⁹ as the approach is generally adequate to cover the large scope of budget support programmes and helped to generate a substantial body of evidence on the effectiveness of budget support, but weak with regard to establishing clear and unambiguous causal links between the different inputs and the outcomes of budget support. Schmitt and Beach (2015) point to the limited value of constructing counterfactuals for causal inference in budget support evaluations of single cases. Overall, scholars and evaluators agree that even more recent evaluations continue to face problems of attribution and that the explanatory power of budget support evaluations is particularly limited for the outcome and impact level.⁵⁰

2.2 Design and methodology

This evaluation synthesis of budget support adopts a theory-based approach and uses the intervention logic laid out in the CEF as the overall analytical framework. The CEF was used to identify the expected outcomes of budget support and to establish the comprehensive code system applied for data collection (see Section 2.2.4). The individual sections in Section 4 build on the CEF to further specify how different elements of budget support were implemented, describe how budget support contributed to results in the different thematic areas,^{xxiii} and identify cross-linkages between different thematic areas.^{xxiv} This evaluation synthesis structures findings on the effectiveness of budget support along seven thematic areas derived from the CEF. The evaluation team concentrated on the issue of effectiveness, which is one out of five DAC criteria for evaluating development assistance.^{xxv} Given that most of

^{xxii} In addition to problems around causation, Dijkstra and de Kemp (2015) point to challenges emerging from practically applying the intervention logic. When adhering to the ownership principle, basic trust in the partner country’s government and some congruence in donors’ and partners’ preferences is required. Another challenge is that budget support programmes are joint efforts by multiple donors. In the joint programme however, individual preferences of donors might deviate.

^{xxiii} The seven thematic areas include: aid effectiveness principles, government expenditure, PFM, corruption, domestic accountability and democratic governance, economic performance and income poverty, service delivery, and non-income poverty.

^{xxiv} The role of theory in synthesis work has been described primarily as an “organizing tool” to structure the synthesized evidence (see Alton-Lee, 2004). Starting from the optimistic programme logic, the present evaluation synthesis will critically assess existing evidence also in the light of competing theories for the individual chapters.

^{xxv} The five criteria are relevance, effectiveness, efficiency, impact, and sustainability (OECD DAC, 1991).

the reviewed evaluations and studies do not explicitly refer to the DAC criteria, this evaluation synthesis focuses on the effectiveness of budget support.

Over the past years, a considerable number of country-based evaluations have generated a substantial amount of evidence on the effectiveness of budget support, which in turn has been summarized in various synthesis studies.⁵¹ These synthesis studies broaden the picture of the effectiveness of budget support in different contexts. However, the studies typically have limited coverage, adopt selective perspectives and do not always review the data in a particularly systematic manner. For example, the criteria for including or excluding primary sources are not always made transparent. Moreover, existing synthesis studies tend to focus on presenting the results without discussing the qualitative heterogeneity and methodological challenges of individual country evaluations in sufficient detail.

Methodological challenges in synthesis work have been the topic of a lively academic debate in the past years. Littell and Maynard (2014) describe three major forms of bias typically found in many research reports and reviews: outcome-reporting bias (in which the reporting of results is influenced by their direction and/or statistical significance), publication bias (by only including published work) and dissemination bias (studies with significant results are published faster, reprinted more often and easier to locate).

In reaction to these challenges, the “systematic review” approach has been proposed as a viable procedure to address multiple biases in traditional research reports. As stated by Saini and Shlonsky (2012), “systematic reviews attempt to decrease the bias of traditional literature reviews by systematically and transparently synthesizing the greatest range of relevant, high-quality studies, published or unpublished, related to a single, pre-specified question.”⁵² Littell and Corcoran (2010) define systematic reviews as “carefully organized, comprehensive, and transparent studies of previous research on a particular topic. Systematic reviews follow written protocols (detailed plans) that specify the central objectives, concepts, and methods in advance.”⁵³ The Campbell Collaboration (2015), an interdisciplinary research

network that disseminates systematic reviews, describes four steps of a systematic review process:

- an explicit search strategy
- clear inclusion/ exclusion criteria
- systematic coding and analysis of included studies
- meta-analysis (where possible).

Systematic reviews typically comprise statistical analyses of high-quality studies, often linked to the use of randomized control trials. Although mostly applied to synthesize and analyse quantitative studies, the approach is increasingly being discussed for synthesizing qualitative studies as well.⁵⁴ Synthesizing the effectiveness of budget support, where counterfactual-based impact evaluations are not available and the existing evaluations and academic literature use different methodologies to analyse the effectiveness of budget support,^{xxvi} requires a more qualitative approach.

This evaluation synthesis adopts a qualitative approach to systematically review the evidence on the effectiveness of budget support. It reviews evaluations of budget support as well as relevant academic and grey literature in order to address the described challenges and limitations of previous syntheses. With a review of 95 sources of three different types, the evaluation synthesis offers broad coverage and the analysis of different perspectives. The reviewed body of evidence is identified based on an explicit search strategy and clear inclusion and exclusion criteria (see 2.2.2), and the analysis includes a quality assessment to control for qualitative differences in the evidence base. Based on systematic coding, the evaluation synthesis includes only those findings that are covered by sufficient evidence (see 2.2.4). Thus, the main focus of this evaluation synthesis is on presenting reliable evidence rather than on summarizing the effects of budget support.

The Campbell Collaboration provides guidelines on when and how to include qualitative studies to augment quantitative systematic reviews. However, no clear guidelines exist on how to conduct a stand-alone qualitative systematic review.⁵⁵ As a pragmatic approach, this evaluation synthesis therefore follows the general standards for systematic reviews,⁵⁶ and the

^{xxvi} See Dijkstra and de Kemp, 2015.

four steps described above, as closely as possible, but does not apply a strict blueprint to the analysis.^{xxvii} The following sections outline how this evaluation synthesis applies the standards of a systematic review process.

2.2.1 A search strategy based on explicit criteria

Inspired by the systematic review methodology, the evaluation team employed specific search strategies for three types of sources.

The first type of source is evaluation reports, including EC/multi-donor evaluations and agency evaluations.^{xxviii} EC/multi-donor evaluations of budget support were compiled by the evaluation team.^{xxix} To ensure completeness, this list has been shared with two external experts in the field of budget support evaluation for comments and further suggestions. The initial list contained 16 items and was then augmented to include 20 EC/multi-donor evaluations of budget support (see list of references). In addition to the 20 EC and multi-donor evaluations, the evaluation synthesis also includes evaluation reports of multilateral development agencies. As those reports are not always available online, the team approached the evaluation units at the World Bank, the African Development Bank (AfDB), the Asian Development Bank (ADB), and the Inter-American Development Bank (IDB). The request resulted in a total of 45 agency evaluations and learning products related to budget support.

The second type of source is academic literature. For this type of source, the evaluation team systematically searched a wide range of online databases^{xxx} for relevant keywords. Publications including the keywords “budget support” OR “budgetary support” in the title, as well as (“budget support” OR “budgetary support”) AND (“effectiveness” OR “effects” OR “results” OR “outcomes” OR “impact” OR “evaluation”) in any field have been considered. The procedure was done in English, German, French and Spanish. The resulting list of academic literature was then checked by two reviewers in order to

eliminate irrelevant titles (i.e. those not dealing with budget support). The initial search resulted in 157 items of academic literature.

The third type of source is grey literature. Given resource and time constraints, the evaluation team limited the search for grey literature by focusing on sources referenced in bibliographies of evaluations and synthesis reports. Including grey literature was considered important in order to minimize publication bias, allowing further triangulation of sources.^{xxxi} Out of the 74 items identified from bibliographies, 59 were available online. Following requests by the evaluation team, the responsible agencies provided three more reports, bringing the total available body of grey literature to 62 items.

Some items from the three groups were reclassified at a later stage of the process. Several sources acquired through the systematic search for academic literature were re-categorized as grey literature if they had not been subject to any kind of peer review process either by academic publishers or research institutes. A complete list of the evaluation reports, academic literature and grey literature included in the review is presented in the list of references.

2.2.2 Relevance criteria for inclusion of sources

To ensure a transparent and systematic procedure, the evaluation synthesis applied predefined relevance criteria for the inclusion or exclusion of sources from the sample. The decision to include or exclude each source was taken independently by two reviewers based on two necessary (“must have”) criteria. Only sources fulfilling both criteria A and B described below were defined as relevant and were included in the evaluation synthesis. Cases of disagreement were discussed and settled in consensus. The two criteria were the following:

^{xxvii} Note that the fourth step (Meta-analysis) proposed by Campbell Collaboration (2015) could not be applied in this evaluation synthesis due to the qualitative character and the variance of applied methodology of the included studies and evaluations.

^{xxviii} EC/multi-donor evaluations include evaluations commissioned by multiple donors and evaluations commissioned by the EU alone but which follow the same methodology as the evaluations commissioned by multiple donors. Agency evaluations include evaluations commissioned by the World Bank and regional development banks. The distinction between EC/multi-donor and agency evaluations was only made in the course of the search strategy.

^{xxix} Some evaluation reports comprise multiple volumes with relevant information (e.g. on methodological issues) provided separately from the main report. When these volumes were not available online, they have been requested from the authors or the commissioning entity.

^{xxx} The databases include: JSTOR, Emerald Insight, Sage Journals, Wiley Online Library, Science Direct, Springer Link, OECD iLibrary, World Bank eLibrary, Google Scholar, Academia.edu, Web of Science, EBSCOhost, WISO, Scopus, Project Muse, Annual Reviews, EconLit.

^{xxxi} Note that triangulation of sources was limited due to different levels of coverage of evidence for different effects of budget support.

Criterion A: The study analyses budget support based on own empirical analysis.

The study is based on own empirical analysis and generates new insights beyond the knowledge from secondary data. Evaluation/studies typically comprise either i) case studies for which a sufficient amount of primary data has been collected or ii) cross-country quantitative studies that go beyond descriptive statistics by conducting (inferential) quantitative analysis of existing datasets.

Criterion B: The study deals with implementation and/or effectiveness of budget support.

It is not required that the study covers the entire spectrum of budget support implementation and questions of effectiveness, but it must deal with at least one of the two aspects. For implementation, important aspects include donor harmonization, transaction costs or predictability of budget support inputs. Important aspects of effectiveness include, among others: poverty reduction, service delivery, corruption, PFM, and domestic accountability.

All EC/multi-donor evaluations and 12 out of 45 agency evaluations passed the relevance criteria. For academic literature, 42 out of 157 items fulfilled the criteria and 21 out of 62 grey literature items passed the criteria.

Quality assessment

The included body of 95 evaluations and studies were assessed regarding their methodological quality, as stated in the reports and supplementary documents. Given the different nature of evaluations vis-à-vis academic and grey literature, two different sets of quality criteria (QC) have been developed and applied by the evaluation team.^{xxxii} Following recent methodological discussions,⁵⁷ quality criteria for evaluations have been defined in several areas: intervention logic and evaluation questions, context, methodology, data collection and analysis, and justification of conclusions. The quality was rated along a 0–3 scale for most quality criteria (see Annex 7.2). For the evaluation reports, the criteria for the quality assessment are presented in Table 1.

^{xxxii} Given the described challenges of evaluating budget support and the bulk of the evidence being produced by evaluations, the focus of the quality assessment was on the evaluation reports. For the academic and grey literature, a reduced set of quality criteria has been applied.

Table 1: Criteria for quality assessment of evaluation reports

Category	Quality criterion
Intervention Logic and Evaluation Questions	QC 1: Formulation of evaluation questions (EQ)
	QC 2: The evaluation answers to the formulated EQs
Context	QC 3: The evaluation describes the economic and political context in the case country
	QC 4: The evaluation describes the context of aid relations in the case country
Methodology	QC 5: The evaluation makes use of an intervention logic (IL)
	QC 6: The evaluation acknowledges problems of attribution and describes how it addresses them in the methodology section
	QC 7: The evaluation reflects on the influence of (changing) context on programme outcomes in the methodology section
	QC 8: The evaluation describes how it triangulates methods and sources in the methodology section
	QC 9: The report describes general limitations of the evaluation in the methodology section
Data Collection and Analysis	QC 10: The evaluation provides information on the collected data (figures, interview partners, document references)
	QC 11: The scope of the empirical inquiry (stakeholder groups)
	QC 12: The scope of the empirical inquiry (number of interviews)
	QC 13: The evaluation provides information on the procedures for data analysis
Justification of Conclusions	QC 14: The evaluation presents its conclusions with reference to methodological limitations
	QC 15: The evaluation makes reference to the information that supports each conclusion
	QC 16: The evaluation identifies and discusses the programme's unintended effects
	QC 17: The evaluation elaborates plausible alternative explanations of the findings
	QC 18: The evaluation justifies why rival explanations were rejected

The evaluation team confined the search for information to particular parts of the evaluation report, focusing on statements made in the methodological section (QC 6, 7, 8, 9) or the conclusion (QC 14, 15). A limitation to this approach is that it is based on the information explicitly given in the evaluation reports and that results depend on the level of transparency of the individual sources. In order to address this limitation, to ensure stakeholder involvement, and to avoid omitting important information, the evaluation team shared information on the quality assessment with the authors of the reports. This information indicated areas (QCs) for which no information was found in the reports and the authors were invited to provide additional information for reassessment.

The quality criteria for academic and grey literature covered the same topics, but included less-specific criteria regarding methodology and data collection and analysis (see Table 2). As a large variety of methods was used across the studies, more general criteria were applied. Two of the quality criteria were changed and reassessed during the quality assessment, as indicated in Table 2, as the questions were too general to be answered for individual studies. The studies were given a score in the range 0–2, where 0 means the QC was not fulfilled, 1 means it was partially fulfilled and 2 means that the study fully adhered to the QC. A more precise rating, as was done in the case of evaluations, was not possible due to the strong differences between the studies whereas the evaluations were easier to compare.

Table 2: Criteria for quality assessment of academic and grey literature

Category	Quality criterion
Research Question	QC1: The study is peer reviewed
	QC2: The study formulates a clear research question (RQ)
	QC3: The study answers the formulated RQ
Methodology	QC4: The study builds on a theoretical framework
	QC5: The study is transparent about the methodological approach. <i>Formerly: The methodological approach is generally adequate to answer the RQ</i>
	QC6: The study provides information on data (collection) and procedures of data analysis.
Context	QC7: The study reflects on the respective context and its influence on the research results. <i>Formerly: The study reflects on the influence of context on research results</i>
Justified Conclusion	QC8: The study establishes a transparent link between the conclusion and the analysed data

The quality assessment provides an overview of the methodological quality of the body of evidence and identifies the strengths and weaknesses of different (groups of) sources (described in more detail in Section 3). The quality assessment finds the overall quality of evaluations and reports to be moderate, but also relatively homogeneous across the sample of evaluations and studies. In view of this result, it would not be justifiable to use the quality score as an additional filter for inclusion or exclusion. Therefore the team decided to discard the initial intention to use the quality assessment as another hurdle to include/exclude sources. All identified evaluations and studies can be considered to be of sufficient methodological quality and are thus included in the evaluation synthesis.

Importantly, differences in quality were taken into account in the course of data analysis and synthesis of results. In order to distinguish between best evidence and second-best evidence, the evaluation team defined a cut-off at 50% of the maximum quality score and conducted a best-evidence robustness check for all identified effects (see 2.2.4). The cut-off at a quality score of 50% of the respective maximum achievable score for evaluations and studies is an arbitrary but pragmatic choice, given that the quality score is fairly evenly distributed across evaluation reports, dividing the body of evidence roughly into half “best” evidence and half “second-best” evidence.

2.2.3 Systematic coding

In line with the requirements set out for systematic reviews, and in order to ensure a maximum level of objectivity, data collection was conducted by developing a comprehensive code system and applying this to the body of included literature. The coding was conducted in two phases. In the first phase, two reviewers engaged in coding, testing and refining the code system and the procedure. This helped to establish and consolidate a common understanding over the use of codes and increased confidence when handling the large dataset. In the second phase, the actual coding was conducted by one evaluator only. Through several rounds of feedback and reassessment among the two evaluators in the course of the process, the reliability of coding results was continuously reviewed. This dialogic approach to data analysis served to ensure a maximum level of objectivity.

Building on the intervention logic of budget support spelled out in the CEF (see Figure 1), a comprehensive code system (see Figure 2) was developed that covers the variety of envisaged effects of budget support. In order to capture additional aspects not covered in the generic intervention logic, the evaluation team went beyond the aspects covered in the CEF and assessed and synthesized evaluation questions actually posed in evaluations of budget support, e.g. regarding gender and corruption. The complete variety of aspects for synthesis was then arranged in a preliminary code system.

This preliminary code system was tested and refined by two reviewers, who applied it to a subset of six sources (evaluations reports and studies) in a first round of coding. As the difference between the codings for macroeconomic effects (induced output level) and economic performance (outcome level) was often not clear-cut in the sources, the codes were merged into one code under the label macroeconomic performance. The same was done in the case of the two codes public spending and allocative efficiency, which were combined under the code composition of public spending. After this a total of 84 codes (including the aggregate codes) was arranged in a comprehensive code system (see Figure 2) and applied to collect data from the full set of relevant sources.

The comprehensive code system was then applied to collect data on effects of budget support. Evaluation reports, academic and grey literature were coded, focusing on the concluding section or, if available, concluding sections within thematic sections. The decision to confine data collection to conclusions has been taken by the evaluation team and discussed with the reference group and peer reviewers. This approach is in line with the interest in actual findings. Moreover, it helped to keep the number of codings at a manageable level and enabled the evaluation team to cover all areas of budget support effects. However, focusing on the conclusion of evaluations comes with the risk of underreporting unintended and negative outcomes that are mentioned in the main text but not in the conclusions.

Figure 2: Comprehensive code system

-
- Context
 - Recipient Context
 - Commitment to Reform
 - Ownership of Poverty Reduction Strategy
 - Capacities to Implement
 - Public Goods Orientation
 - Donor Context
 - Donor Group Tensions
 - Political Economy in Donor Country
 - Capacities at Country Level
 - Aid Relations Context
 - Aid Dependency
 - Budget Support Dependency
 - Joint Assistance Framework
 - Size of Donor Group
 - Lead in Donor Group
 - Tensions in Aid Relations
 - Initial Context
 - Input
 - Input Codes
 - Budget Support Programme
 - Funds
 - Policy Dialogue
 - Conditionality
 - Technical Assistance/ Capacity Development
 - Combine Paris Principles
 - Ownership
 - Harmonization
 - Alignment
 - Predictability
 - Transaction Costs
 - Outcome
 - Direct Outputs
 - On Budget Total
 - Ownership Total
 - Harmonization Total
 - Alignment Total
 - Predictability Total
 - Transaction Costs Total
 - Induced Outputs
 - Macro
 - Fiscal Space / Level of Public Spending
 - Domestic Revenue Mobilization
 - Public Financial Management
 - Budget Management
 - Budget Transparency
 - Composition of Public Spending
 - Reform / Policy Process
 - Allocative Efficiency
 - Domestic Accountability
 - Role of Parliament
 - Civil Society Organizations
 - Supreme Audit Institutions
 - Anti-Corruption Agency
 - Corruption Policies
 - Corruption Legislation
 - Corruption Level
 - Outcome and Impact
 - Admin for Service Delivery
 - Service Delivery
 - Economic Performance
 - Non-Income Poverty
 - Income Poverty
 - Income Distribution
 - Other Intended Outcomes/Impacts
 - Unintended Outcomes
 - General Impacts and Cross Cutting Issues
 - Gender
 - Governance
 - Democratization
 - Decentralization
 - Private Sector Development
 - Liberalization
 - Peace and Security
 - Sector
 - Agriculture
 - Education
 - Health
 - Infrastructure
 - Other Sector
 - Success / Failure Factors
 - Success Factor
 - Failure Factor

All 95 included sources were systematically coded, which resulted in 5856 codings (i.e. text sections). The coding process was conducted using MAXQDA®, a software package for qualitative data analysis. Text segments were coded using code combinations to specify the particular outcome area (e.g. PFM, domestic accountability, service delivery etc.) as well as the specific inputs (i.e. the set of budget support inputs – funds, policy dialogue, conditionality, TA/CD)^{xxxiii} that affected the changes in outcome. In order to document the direction of individual effects, labels were assigned to the coded text segments rating effects as positive,^{xxxiv} negative, absent, improving, or deteriorating. The labels improving and deteriorating, compared to positive and negative, were assigned when effects changed in the course of budget support disbursements. The categories for the effect direction were established during the first round of coding. It is important to stress that the classification of an effect's direction depends on the strength of wording that describes the effect and therefore is not an exact and objective measure of strength of the respective effect, but serves as a proxy. As a consequence, effects are not comparable across different studies and evaluations conducted by different scholars or consultants. Notwithstanding these limitations, the effect size labels provide important orientation with regard to the overall direction (positive, negative, absent) for the entire body of evidence (Section 3) and the individual effects (see Section 4).

Context factors were included for data collection, and the evaluation team coded different dimensions of context (country context, aid relations context, donor context) as well as success and failure factors. Unfortunately, however, the reports and studies provide context information mainly at a general level without linking individual context factors explicitly to the specific effects of budget support. The influence of context on budget support effectiveness is therefore described to a varying extent for the different effects in Section 4. Given these limitations in the evidence base, synthesis results on context factors could only be established for a subset of the

effects, based on the identified success and failure factors described in the sources.

2.2.4 Data analysis and reporting

In order to provide a structure for reporting, the evaluation team clustered related codes along the lines of seven thematic areas: aid effectiveness principles, government expenditure, PFM, corruption, domestic accountability and democratic governance, economic performance and income poverty, and service delivery and non-income poverty. These clusters combine multiple codes from a common thematic area (i.e. the cluster PFM contains the codings on overall PFM, budget management and budget transparency) and include various effects across different levels of the intervention logic.

The evaluation team conducted a quantitative analysis of the frequency of codings and a qualitative analysis of the effects for the seven thematic areas. The quantitative analysis built on descriptive statistics of the frequency of codings, sources reporting the effect, and effect directions. It provided important information on how the different thematic areas are covered by the body of evidence. The analysis identified those outcomes that are well covered by evidence and those that are not, or are insignificantly covered. Both presence and absence of budget support effects are treated as important findings of this evaluation synthesis. The aggregate number of sources and codings, as well as the direction of effects (positive, negative or absent) compiled as part of the quantitative analysis are further used to underpin the qualitative findings.^{xxxv}

After finalizing the process of coding, the evaluation team conducted a qualitative analysis. To this end, the dataset was exported into a spreadsheet listing all text segments, with additional information such as effect weight, country, quality assessment (best evidence, second-best evidence), and input codes coded to the same text segment (budget support programme, funds, policy dialogue, conditionality, TA/CD). For each thematic area, the qualitative analysis followed a

^{xxxiii} Note that for most text segments, no specific input of budget support has been identified. Therefore, most effects have been linked to the budget support programme in general using the code "Budget Support Programme".

^{xxxiv} Positive effects were initially rated more fine-grained as weak, moderate, strong, which helped to structure the evidence in the qualitative data analysis.

^{xxxv} Note that the number of sources supporting individual findings from the qualitative analysis can diverge from the overall amount of statements identified in the quantitative analysis. The quantitative analysis draws on the number of text segments per code and effect direction thereby accounting for all segments as coded. The qualitative analysis comprised a more selective approach in which the evaluator decides upon the use of individual text segment while doing the qualitative analysis.

standardized procedure involving three main steps. First, the evaluators carefully read all codings for the particular thematic section. Second, the evaluators used paraphrases and keywords to further structure the evidence. Third, by sorting and grouping codings in different ways (by country, by effect weight, by quality, by paraphrase content and by input code), the evaluators were able to identify scope and patterns of identified effects. Drawing together the results from the qualitative and quantitative analysis allowed the evaluation team to ensure internal transparency of the evidence base and to carve out a comprehensive picture of the evidence base on budget support in relation to the programme theory underlying the generic intervention logic of the instrument (see Section 4).

In order to assess the reliability of findings, the evaluation team applied multiple measures in the data analysis and reporting phase. First, the team identified reported effects with a questionable attribution to budget support which have not been considered in the qualitative analysis. This has been the case in particular regarding effects of budget support on outcome and impact level.

Second, the data analysis includes a robustness check. Based on the quality assessment and its resulting categorization of sources in best evidence and second-best evidence, the robustness check examines the composition of sources from the two categories. Findings that are backed up by sources in the “best evidence” category are considered robust. Findings from best and second-best evidence pointing in the same direction are also considered robust. In cases where statements are covered by both best and second-best evidence and contain contradictory effects, the lack of robustness is noted and discussed. Note that in all effects presented in Section 4, findings were covered by either only best or both best and second-best evidence pointing in the same direction, showing the robustness of findings.

Third, effects are only reported in this evaluation synthesis when they are covered in at least ten sources. This is obviously an arbitrary choice of what represents a “sufficient” evidence base. On the one hand, one could argue that 10 out of 95 sources seem to be a low benchmark. On the other hand, it reflects a rather conservative approach given the variety of expected effects of budget support: As the majority of studies focus on a subset of effects and the total number of sources covering one specific effect is way lower than 95, the coverage of an effect in 10 or more sources is only achieved in the case of some frequently reported effects. Thus, the chosen cut-off serves the purpose of this evaluation synthesis to identify the secured effects of the aid instrument. In order to make the implications of this choice as transparent as possible for the reader and to allow for stricter or more lenient interpretations of the evidence base, the text indicates for all findings in how many sources each finding was stated.



3.

BODY OF EVIDENCE

Table 3: Characteristics of evaluations reports

Document group	Institution	Published	Research Period	Covered Countries	Design
Evaluations (32)	EC (9)	2011–2016	1996–2015	Burundi, Mali, Morocco, Mozambique, Sierra Leone, South Africa, Tanzania, Tunisia, Uganda	Case studies
	Consortium IDD, Mokoro, Ecorys, DRN and NCG (7)	2006	1994–2004	Burkina Faso, Malawi, Mozambique, Nicaragua, Rwanda, Uganda, Vietnam	Case studies
	Others (4)	2005–2012	1995–2010	Ghana, Nicaragua, Tanzania, Zambia	Case studies
	World Bank (9)	2008–2015	2007–2014	Burkina Faso, Dominican Republic, Ethiopia, Mozambique, Santa Lucia, Tanzania, Uganda, Vietnam	Case studies
	AfDB (2)	2011	1999–2009	Sierra Leone, Tanzania	Case studies
	IDB (1)	2015	2014	No country cases	Technical note

This section provides information on the body of evidence. It describes general characteristics of the sources used in this evaluation synthesis as well as the results and interpretation of the quality assessment. Second, it quantitatively describes the evidence, identifying the outcomes that are well covered by evidence and those that are not, or are insignificantly covered. Whereas the aggregate numbers of sources and codings are part of this section, the quantitative findings relating to specific outcomes are described in Section 4.

3.1 Characteristics of the sources of evidence

This evaluation synthesis assesses the evidence on budget support effects on multiple outcomes in 95 documents. Of the 95 documents, 32 are evaluations, 42 are academic literature, and 21 are grey literature.

The 32 evaluations were all undertaken by the EC, a consortium led by the International Development Department of the University of Birmingham (IDD), the World Bank, the AfDB, or the IDB, and published between 2005 and 2016. The evaluations altogether cover an evaluation period from 1994 to 2015^{xxxvi} and are, with one exception, based on a case-study design, focusing on one country each (see Table 3 for information on each category).

^{xxxvi} The latest WB evaluations, published in 2016, are not included in this evaluation synthesis as they were only published after data collection for this synthesis was complete.

Table 4: Characteristics of academic and grey literature

Document Group	Institutions	Document type	Published	Research Period	Covered Countries	Design
Academic Literature (42)	e.g. DIE, EP, IMF, IOB, ODI, UNU-WIDER, WB, WHO	Discussion papers (22)	2005–2015	1992–2012	43 countries ^{xxxvii}	Case studies (9) Comparative case studies (5) Large n (8)
	e.g. IOB	Journal Articles (8)	2008–2015	2000–2011	6 countries ^{xxxviii}	Case studies (4) Comparative case studies (1) Large n (3)
	e.g. IDD, IOB, ODI	Reports (12)	2000–2015	1994–2014	13 countries ^{xxxix}	Case studies (6) Comparative case studies (2) Large n (3) Synthesis (1)
Grey Literature (21)	e.g. ecdpm, ISS, KAS, Oxfam, PPC, U4	Discussion papers (10)	2004–2013	1995–2011	12 countries ^{xl}	Case studies (6) Comparative case studies (1) Large n (3)
	e.g. EU, OPM/ODI	Reports (11)	2003–2014	1992–2014	12 countries ^{xli}	Case studies (9) Synthesis (2)

The academic literature on budget support can broadly be categorized into discussion papers, journal articles and reports. They are published between 2003 and 2015. The discussion papers cover a sum of 46 countries, the journal articles six countries and the reports 15 countries. Altogether, the research period ranges from 1992 to 2014. In terms of design, academic and grey literature use either a case-study design, a comparative case-study design or a large-n design. See Table 4 for more information on each category.

3.2 Reliability of evidence: methodological quality

Quality assessment of evaluation reports

The final results of the quality assessment for evaluations are displayed in Table 5. Each row indicates the quality assessment for one source. As described in Section 2.2.2 and in Annex 7.2, the quality of each report was assessed on a scale of 0–3 points for most of the QC.^{xliii} The maximum possible score is 49 points.

All the assessed evaluation reports are placed in a middle range (i.e. second and third quartile), with quality scores from 15 to 34. Not one evaluation report comes close to reaching the maximum score against standard quality criteria for evaluations. This result arguably reflects the important methodological challenges in evaluating budget support. At the same time, the quality assessment strongly suggests that the overall quality of evaluation work on budget support is sufficiently strong to synthesize conclusions from the overall body of evidence.

To distinguish between evidence backed by stronger quality evaluations and those of lesser quality, based on the applied criteria, the sample was divided into best evidence (top half of the table) and second-best evidence (bottom half of the table). Table 5 and Table 6 indicate the division into best and second-best evidence by the grey line. 17 out of 32 evaluations are thus categorized as best evidence.

^{xxxvii} Angola, Benin, Botswana, Burkina Faso, Burundi, Cameroon, Cape Verde, Central African Republic, Chad, Congo, Comoros, Côte d'Ivoire, Djibouti, Dominican Republic, DRC, Equatorial Guinea, Eritrea, Ethiopia, Gabon, Gambia, Ghana, Guinea, Guinea Bissau, Kenya, Lesotho, Liberia, Madagascar, Malawi, Mali, Mauritania, Mozambique, Namibia, Niger, Nigeria, Rwanda, Senegal, Sierra Leone, Sudan, Tanzania, Togo, Uganda, Zambia, Zimbabwe.

^{xxxviii} Ethiopia, Ghana, Nicaragua, Rwanda, Tanzania, and Zambia.

^{xxxix} Burkina Faso, Ghana, Malawi, Mali, Mozambique, Nicaragua, Rwanda, Sierra Leone, South Sudan, Tanzania, Uganda, Vietnam, and Zambia.

^{xl} Afghanistan, Burundi, DRC, Ethiopia, Ghana, Malawi, Mozambique, Nicaragua, Palestine, Tanzania, Uganda, and Vietnam.

^{xli} Andhra Pradesh, Benin, Burkina Faso, Georgia, Ghana, Moldova, Mozambique, Nicaragua, Tanzania, Uganda, Ukraine, and Zambia.

^{xliii} QC8, QC14, QC17 and QC18 use a different scale, see Annex 2.

Table 5: Quality assessment of evaluation reports

Document group	QC 1	QC 2	QC 3	QC 4	QC 5	QC 6	QC 7	QC 8	QC 9	QC 10	QC 11	QC 12	QC 13	QC 14	QC 15	QC 16	QC 17	QC 18	Sum
Evaluations	3	3	3	3	3	2	2	1	1	2	3	3	2	0	1	1	1	0	34
	3	3	1	3	1	2	1	1	1	3	3	3	1	0	3	1	2	0	32
	3	3	3	2	2	0	0	1	3	3	3	3	2	0	2	0	1	0	31
	3	3	3	1	3	0	0	1	1	3	3	3	1	0	3	2	0	0	30
	3	3	1	2	1	2	1	1	1	3	3	3	1	0	3	1	1	0	30
	3	3	1	3	1	2	1	1	1	3	3	3	1	0	3	1	0	0	30
	3	3	1	2	1	2	1	1	1	3	3	3	1	0	3	0	1	0	29
	3	3	1	3	1	2	1	1	1	3	3	2	1	0	3	0	1	0	29
	3	3	1	3	1	2	1	1	1	1	3	3	1	0	3	0	1	1	29
	3	3	3	2	2	0	0	1	0	2	3	3	2	0	2	1	1	0	28
	3	3	3	1	2	0	0	1	1	2	3	3	0	0	3	2	1	0	28
	3	3	1	2	1	2	1	1	1	2	3	3	1	0	3	0	1	0	28
	2	3	3	2	1	1	2	1	2	2	3	2	0	0	2	0	1	0	27
	2	3	2	1	0	1	2	2	2	2	3	3	2	0	1	1	0	0	27
	3	1	1	1	3	0	1	1	2	3	3	3	1	0	3	0	1	0	27
	3	3	3	1	2	0	0	0	0	2	3	3	0	0	3	2	1	0	26
	2	3	2	3	0	0	2	0	2	2	3	3	0	1	2	0	0	0	25
	2	3	2	1	0	1	1	0	1	1	3	3	0	1	2	1	1	0	23
	2	3	2	1	1	1	0	0	1	2	3	3	0	1	2	0	1	0	23
	3	3	1	1	2	0	0	1	2	3			0	0	3	1	2	1	23
	2	1	2	1	1	0	0	1	3	2	3		1		0	2	2	1	22
	2	2	1	1	0	1	2	0	1	2	2		2	1	2	2	0	0	21
	3	3	1	0	1	0	0	2	1	2	3	3	0	0	0	0	1	1	21
	3	3	1	1	1	1	0	0	0	2	3	3	0	0	0	1	1	1	21
	2	3	1	1	1	0	0	1	0	2	3	3	0	1	2	0	0	0	20
	2	3	2	1	0	0	2	0	0	1	3	3	0	0	2	0	0	0	19
	2	3	2	1	0	0	1	0	0	1	3	2	0	0	2	1	1	0	19
	2	3	1	1	0	1	1	0	1	1	3		0	1	2	1	0	0	18
	3	2	2	0	2	0	0	1	2	1			1	0	3	0	1	0	18
	2	2	2	0	0	1	0	0	1	1	3		0	1	2	2	0	0	17
	2	3	2	0	0	1	1	0	1	1	3		0	1	2	0	0	0	17
	2	3	2	0	0	0	1	0	0	1	2	1	0	0	2	1	0	0	15
Total score per QC	82	89	57	45	34	25	25	22	35	64	88	70	21	8	69	24	24	5	

The criteria 1–8 for the quality assessment of evaluation reports are listed in Table 1. The quality was rated along a 0–3 scale for most quality criteria. The rating scale for each criterion is presented in Annex 7.2. Depending on the given rating, the cells in this table were colour-coded, 0 in light yellow, 1 in yellow, 2 in light blue and 3 in blue. Empty cells indicate that the QC could not be rated due to unspecified information in the source (e.g. the source states that interviews were conducted, but contains no statement on how many and with whom). In contrast, missing information was rated with zero. The ratings for each quality criteria were added to build the total score per QC (last row in table). To recognize patterns of highest and lowest scores, the total score was also colour-coded. The high rankings were coded blue, dark blue for the two highest rankings and lighter shades of blue for high rankings, while low rankings were coded orange, with dark orange for the two lowest rankings and lighter orange for low rankings.

With regard to the individual quality criteria, the three criteria that scored highest among all evaluations are QC2, QC11, and QC1 (in order of highest score – see dark blue colour coding in Table 5). The high score for the formulation of an evaluation question (QC1) and answering this question (QC2) indicates that the evaluations have a clear focus and manage to provide findings on this question. QC11 refers to the scope of the empirical inquiry of different stakeholder groups, and the high score implies that almost all of the budget support evaluations cover both perspectives in data collection (donor and partner country). Other criteria that scored above average are QC12, QC15, and QC10 (in order of highest score – see blue colour coding in Table 5). This shows that the evaluations rely on a considerable amount of data collected in the form of interviews (QC12), usually make reference to the information that supports each conclusion (QC15), and provide information on the collected data (QC 10). The high ratings on data collection for all stakeholder groups, based on a large number of interviews, are a positive indication of the quality of the data on which the evaluations are based.

The criteria that scored lowest among all the QCs are QC18

and QC14 (see orange colour coding).^{xliii} Very few of the evaluations mention the limitations of the evaluation design in the conclusion (QC14) and also few explain why rival explanations are rejected (QC18). Other criteria that scored low are QC13, QC8, QC16, and QC17, indicating that few evaluations mention a triangulation of methods (QC8) and identify possible alternative explanations (QC17).^{xliiv} Also, few mention the information on the procedures for data analysis (QC13) and the discussion of unintended effects (QC16). It is noteworthy that quality criteria on limitations or unintended effects of the evaluations had particularly low scores (QC16, QC17, QC18), pointing to the risk of a positive bias in budget support evaluations. To address this bias, findings from evaluations were triangulated with findings from academic and grey literature as far as findings were available in both document types.

Also, the criterion QC6 (the evaluation acknowledges problems of attribution and describes how it addresses them in the methodology section) scored low. This illustrates the unresolved methodological challenges in budget support evaluations regarding the attribution of effects at outcome and impact level to budget support. The problematic attribution at this level is also a central finding of the qualitative analysis presented in Section 4.

Quality assessment of academic and grey literature

The results of the quality assessment of academic and grey literature are reported in Table 6. The quality scores range from 0 to 15, with a maximum possible score of 16. The assessed studies thus cover the entire range of quality scores. The grey line in the middle of the table again indicates the division between best evidence (top half of the table) and second-best evidence (bottom half of the table). Empty cells indicate that the QC could not be rated due to unspecified information in the source.

Overall, the results of the quality assessment of academic and grey literature show a large spread across the entire range of quality scores. The academic literature scored better than grey literature on the applied QC (see Table 6), most likely as a

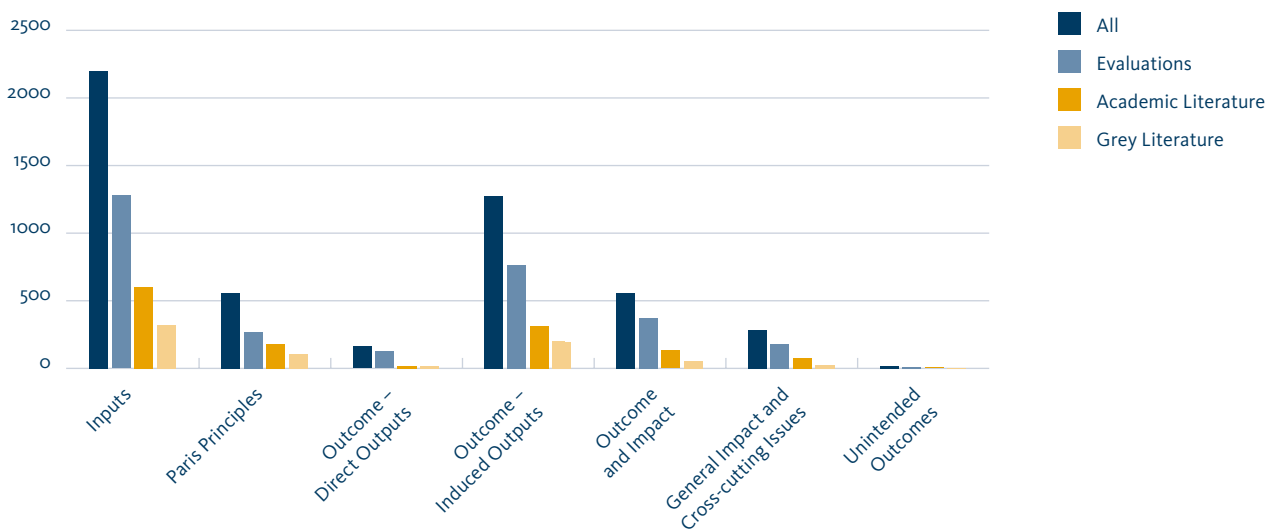
^{xliii} While the other QCs were rated on a scale from 0–3 points, QC14, 17 and 18 were rated along a different scale. QC 14 and QC17 were rated from 0–2 and QC 18 was rated 0 or 1. Due to the higher specificity, these criteria were rated on less detailed scales. However, even if the maximum score had been 3 points, very few of the reports actually obtained the maximum ranking for QC 14 and QC 18, which means these QC would still score very low.

^{xliiv} This low score might also be influenced by the fact that these QC were rated along a 0–2 ranking system and the score could have been higher if a 0–3 ranking had been used.

Table 6: Quality assessment of academic and grey literature

Document type	QC 1	QC 2	QC 3	QC 4	QC 5	QC 6	QC 7	QC 8	Sum
Academic literature	1	2	2	2	2	2	2	2	15
Academic literature	1	2	2	2	2	2	2	2	15
Academic literature	2	2	2	2	2	2	1	2	15
Academic literature	1	2	2	2	2	2	2	2	15
Academic literature	1	2	2	2	2	2	2	2	15
Academic literature	2	1	2	2	2	1	2	2	14
Academic literature	1	1	2	2	2	2	2	2	14
Academic literature	2	2	2	2	2	1	2	1	14
Academic literature	1	2	2	2	2	2	1	2	14
Academic literature	1	2	2	2	1	2	2	2	14
Academic literature	1	2	2	2	2	2	2	1	14
Academic literature	1	1	2	2	2	1	2	2	13
Academic literature	1	2	2	2	2	2	0	2	13
Academic literature	1	2	2	2	2	2	1	1	13
Academic literature	2	2	2	2	2	0	2	0	12
Grey literature	0	2	2	2	2	1	1	2	12
Academic literature	2	2	2	2	2	1	1	0	12
Academic literature	1	2	2	1	2	2	0	2	12
Academic literature	2	2	2	2	1	1	2	0	12
Grey literature	0	2	2	0	2	1	2	2	11
Academic literature	1	2	2	0	2	1	1	2	11
Academic literature	1	2	2	2	1	1	1	1	11
Grey literature	0	1	2	2	2	2	0	2	11
Academic literature	1	1	2	2	2	2	0	1	11
Academic literature	1	0		2	2	2	2	2	11
Grey literature	0	1	2	0	2	1	2	2	10
Academic literature	1	2	2	0	1	1	1	2	10
Academic literature	1	2	2	2	2	0	1	0	10
Academic literature	1	2	2	1	0	1	2	1	10
Grey literature	0	2	0	2	2	2	0	2	10
Grey literature	0	1	2	2	0	0	2	2	9
Grey literature	0	1	2	2	1	0	2	1	9
Academic literature	1	1	2	0	0	2	1	2	9

Document type	QC 1	QC 2	QC 3	QC 4	QC 5	QC 6	QC 7	QC 8	Sum
Academic literature	1	2	2	1	0	1	1	1	9
Academic literature	1	2	2	0	1	1	2	0	9
Academic literature	1	2	2	1	1	1	1	0	9
Academic literature	1	2	2	1	1	1	1	0	9
Grey literature	0	1	2	0	1	1	1	2	8
Academic literature	1	1	2	0	0	1	2	1	8
Academic literature	1	2	2	1	1	0	1	0	8
Grey literature	0	1	1	2	0	1	2	1	8
Academic literature	1	2	2	1	0	1	1	0	8
Academic literature	2	1	0	2	0	0	2	0	7
Academic literature	1	2	1	0	0	1	2	0	7
Grey literature	0	2	0	0	2	2	1	0	7
Grey literature	0	1	2	0	0	0	2	1	6
Academic literature	1	1	2	0	0	0	2	0	6
Academic literature	1	2	2	0	0	0	1	0	6
Grey literature	0	1	2	0	1	1	0	1	6
Grey literature	0	1	2	0	0	1	1	1	6
Academic literature	1	0		0	2	2	0	1	6
Grey literature	0	1	1	1	0	0	1	1	5
Academic literature	1	1	2	0	0	0	1	0	5
Academic literature	2	1	1	0	0	0	1	0	5
Academic literature	1	1	0	1	0	1	1	0	5
Grey literature	0	0		0	1	1	1	2	5
Grey literature	0	1	1	0	0	0	0	2	4
Grey literature	0	1	1	0	1	1	0	0	4
Academic literature	1	0		0	0	0	1	1	3
Grey literature	0	0		0	2	1	0	0	3
Grey literature	0	1	0	0	0	0	0	0	1
Grey literature	0	0		0	1	0	0	0	1
Grey literature	0	0		0	0	0	0	0	0
Total score per QC	50	88	96	65	70	65	75	66	

Figure 3: Frequency of codings by analytical level and document type

result of more stringent quality assurance processes for peer-reviewed publications.

With regard to the individual criteria, the studies score highest on QC3 and QC2 (in order of highest score – see blue and dark blue colour coding in Table 6), rating the inclusion of a research question (QC2) and the answers to this question (QC3) in the studies. QC1 (the study is peer reviewed) reached the lowest score, as all grey literature was rated zero for this QC (as stated in section 2.2.1, the existence of a peer review process was decisive in the classification as academic or grey literature). No particular patterns can be distinguished in the distribution of scores across the remaining QC, and thus no further conclusions can be drawn on general quality characteristics of this body of evidence.

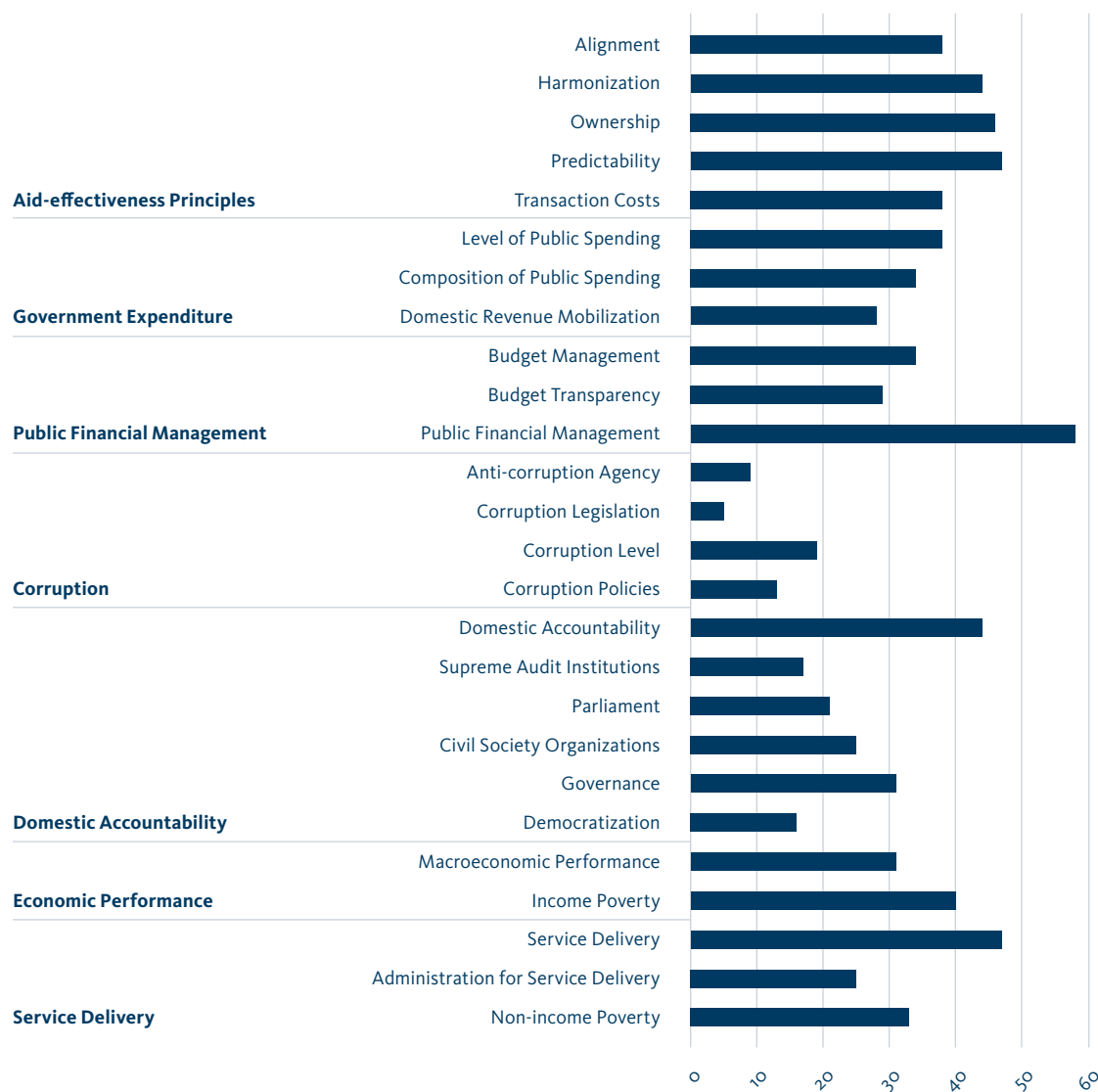
The criteria 1-8 for the quality assessment of academic and grey literature are listed in Table 2. The studies were given a score in the range 0–2, where 0 means the QC was not fulfilled, 1 means it was partially fulfilled and 2 means that the study fully adhered to the QC. Depending on the given rating, the cells in this table were colour-coded, 0 in light yellow, 1 in yellow, and 2 in light blue. Empty cells indicate that the QC could not be rated due to unspecified information in the

source (e.g. the source states that interviews were conducted, but contains no statement on how many and with whom). In contrast, missing information was rated with zero. The ratings for each quality criteria were added to build the total score per QC (last row in table). To recognize patterns of highest and lowest scores, the total score was also colour-coded. The high rankings were coded blue, dark blue for the highest ranking and middle blue for the second highest ranking, while the lowest ranking was coded orange. White cells mark the middle categories, being neither particularly high nor low.

3.3 Aggregate number of sources and codings

In total, 5856 codings were applied to the 95 sources. Each text section on the effect of a budget support input on an outcome was coded. The budget support input either refers to the overall programme or specifically to funds, policy dialogue, conditionality, or TA/CD. The outcomes consist of the compliance or non-compliance with the described aid-effectiveness principles, direct outputs, induced outputs, outcomes and impacts (see Figure 3). In addition, text sections on cross-cutting issues (e.g. gender, peace and security) and unintended outcomes were coded in the sources. Figure 3 shows the total number of codings per analytical level and document type.

Figure 4: Aggregate number of sources per outcome



The highest number of codings relates to inputs (summarizing all budget support inputs: the overall budget support programme, budget support funds, policy dialogue, conditionality and TA/CD) (see also Section 2). Other codes with high coverage are the implementation in accordance with aid-effectiveness principles,^{xiv} induced outputs, outcomes and impacts, each with more than 500 codings for the aggregate of all document types. Among the document types, the highest number of codings is found in evaluations. The number of

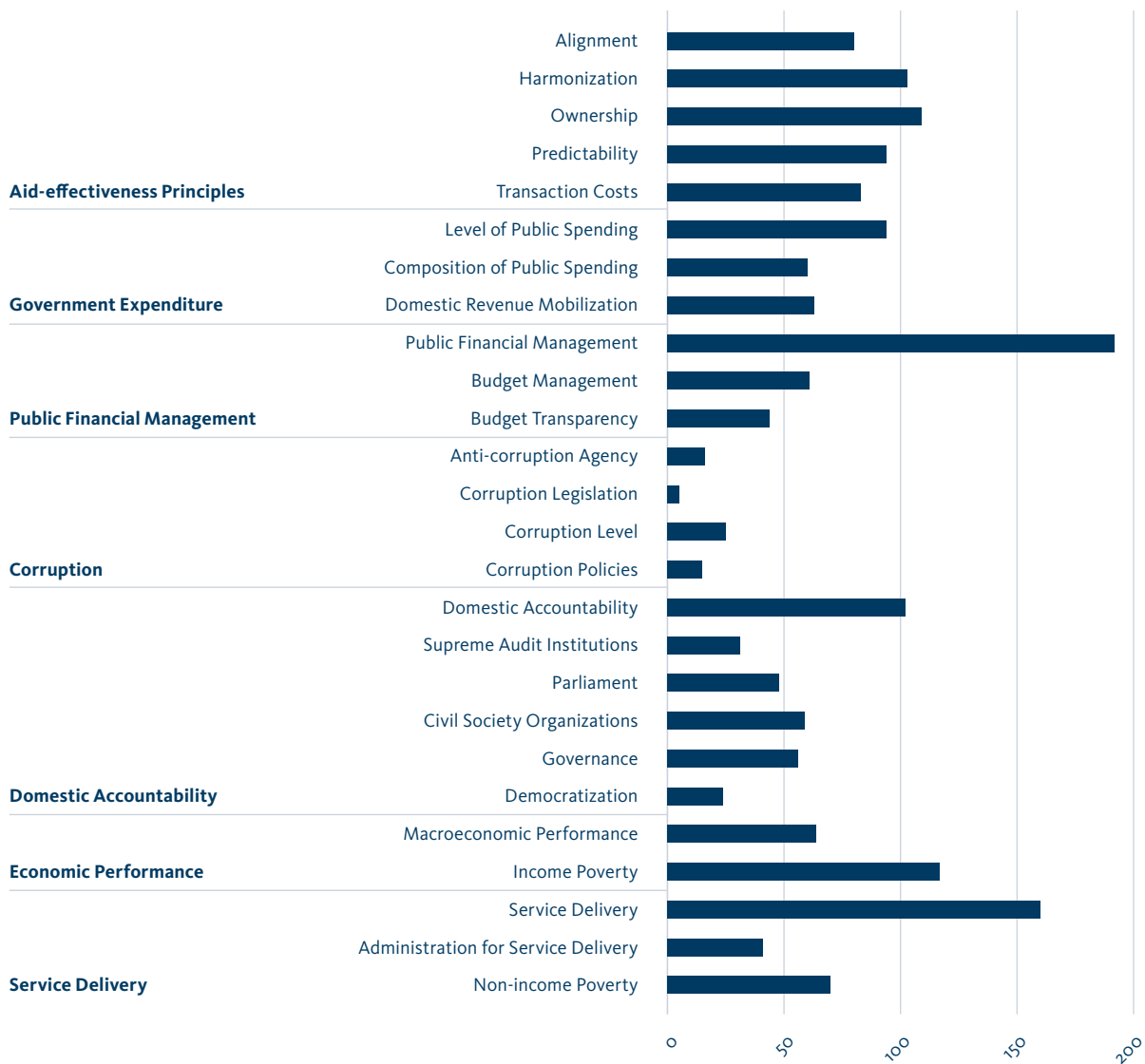
codings is lower for direct outputs, as the difference between the implementation of budget support following aid-effectiveness principles and direct outputs of budget support programmes has not always been clearly stated in the sources. Due to this fact, the codings on the aid-effectiveness principles already include some information on direct outputs. The large number of codings on outcome and impact level was critically assessed regarding the convincing attribution of the effects to the budget support programme. In cases where the attribution

^{xiv} Including the principles of alignment, harmonization and ownership, predictability and reduced transaction costs.

of the effect to the budget support programme was insufficiently traceable, the effect was not considered in the qualitative analysis.

The different codes for outcomes and impact were clustered into the broader topics “government expenditure”, “corruption”, “public financial management”, “domestic accountability”, “service delivery” and “economic development”. Figure 4 shows the sum of all sources^{xlvi} and Figure 5 the sum of all

codings on budget support effects on the respective outcomes. The best coverage, both in terms of sources and codings, is available for budget support effects on PFM, with over 220 codings in 64 out of 95 sources. Also well covered by evidence are budget support effects on service delivery, domestic accountability, income poverty, level of public spending, and three of the aid-effectiveness principles (harmonization, ownership, and predictability), which are mentioned in more than 40 out of 95 sources (see Figure 4).



^{xlvi} Every report is only counted once per outcome.

Table 7: Aggregated effects of specific budget support inputs on all outputs, outcomes and impacts

Aggregate	Direct outputs, induced outputs, outcomes and impact combined	Inputs	Total	BS Programme	Funds	Policy dialogue	Conditionality	TA / CD
		# Sources	85	80	57	42	59	37
Total effect codings	2098	1255	358	181	182	122		
Positive effect	1551	923	279	145	106	98		
No effect	353	226	44	23	37	23		
Negative effect	134	67	22	8	36	1		
Improving effect	25	15	6	2	2	0		
Deteriorating effect	35	24	7	3	1	0		

Low coverage is found for budget support effects on corruption legislation and anti-corruption agencies. These effects are covered in fewer than ten sources, with a maximum of 20 codings per topic. All other codings lie in between, i.e. they are included in fewer than 40, but more than ten sources.

Figure 5 shows the sum of all codings. Note that several text sections on an outcome can be coded per report. The highest number of text sections is coded for PFM (nearly 190) and the lowest are coded for corruption policies, corruption legislation, and anti-corruption agencies (each fewer than 20).

3.4 Strength of budget support effects on output and outcome level

In addition to the aggregate numbers of sources and codings, the direction of identified budget support effects was analysed. The quantitative analysis assesses whether budget support led to positive, negative, or absent effects on the central outcomes. Note that the last category “absent effect” is different from the case where a particular effect is not covered by the sources. It refers to a situation when an expected effect is sufficiently covered in the body of evidence, but no effect (positive or negative) could be found, either because there was no effect or because it could not be detected with the selected method.

For effects on all levels (direct output, induced output, outcome, and impact level), 2098 codings were assigned in 85 sources (see Table 7). Among those 2098 codings, 1551 text sections were described as a positive effect, while 134 were

rated negative. 353 text sections mention the absence of any effect or the absence of evidence for effects of budget support on different outcomes.

Table 7 illustrates in how many sources and codings budget support effects on a specific code are reported. It differentiates between text sections that report effects of budget support programmes in general, which can be found in the column “BS programme” and text sections that report an effect of a specific input (funding, policy dialogue, conditionality, or TA/CD), as seen in the last four columns in Table 7. The column “total” indicates the total number of codings on the GBS programmes and the specific inputs together. The number of codings is further divided by the direction of the effect (weak positive, moderate positive, strong positive, no, negative, improving, or deteriorating). The orange bars indicate the number of codings for each effect direction compared to the total number of codings on the combination of code and input. This explanation also refers to all other graphs of this type in Section 4.

As can be seen in Table 7, the text sections mostly refer to effects of the budget support programmes in general, while the differentiation between financial inputs (funding) and non-financial inputs (policy dialogue, conditionality, TA/CD) is made in fewer cases. The largest number of codings are found on positive effects, and the fewest codings are found on improving and deteriorating effects. Of the specific inputs, the effect of funding has the most codings and the effect of TA/CD has the fewest codings.



4.

EVIDENCE ON BUDGET SUPPORT

This section synthesizes the evidence on effects of budget support in different thematic areas including aid-effectiveness principles, government expenditure, PFM, corruption, domestic accountability, economic performance and income poverty, and service delivery and non-income poverty. The results are drawn from the qualitative and quantitative data analysis.^{xlvii}

Each sub-section in this section follows the same structure. The analysis is divided into three parts: first, a description of the expected budget support effects as laid out in the intervention logic recognized by the OECD DAC, second, a description of the coverage of evidence for each specific budget support effect, and third, a description of the empirical evidence for the respective effect. The analysis of the evidence sub-divides coverage into the attribution of generic effects to the entire budget support programme, and the attribution of input-specific effects to individual budget support funding, policy dialogue, conditionality and TA/CD.

The analysis of the empirical evidence differentiates between “empirically established effects”, for which there is sufficiently strong evidence in the sources (the effect is mentioned in 10 or more reports), and those expected effects for which no generalizable conclusions can be drawn due to insufficient coverage in the sources. As one of the goals of this evaluation is to inform the political debate with reliable evidence, the findings on insufficient evidence are considered equally important as the established evidence on budget support effects. The coverage indicates in which topics more in-depth work on effects and underlying causal mechanisms is necessary.

Among the empirically established effects, the evidence-base for an effect is deemed *sufficient* when the effect is covered in 10 – 34 reports, as *strong* when the effect is covered in 35 – 49 reports and as *very strong* when the effect is covered in 50 – 95 reports. As an additional step, the quality assessment (see Section 2.2), with its resulting categorization of sources into best evidence and second-best evidence is used as a best-evidence robustness check for all effects.

The coverage of evidence for all budget support effects at the level of direct outputs, induced outputs, outcomes and impact, as laid out in the CEF, is illustrated in Table 8. All reported effect directions are summarized in Table 9. A detailed description of the findings is provided in sections 4.1 to 4.6.

Attribution of the effects to budget support on a generic level – insufficient coverage of the effects of specific budget support inputs

A large number of sources attribute observed effects to budget support programmes on a generic level. In most cases, the evidence base is insufficient to link the observed effects to specific budget support inputs (see Table 8). This lack of input-specific attribution and in-depth work on the effects and underlying causal mechanisms of budget support limits the lessons that can be learned from the studies on and evaluations of budget support. However, on the generic level, there is a sufficient evidence base on budget support effects to analyse the effect directions (see Table 9). Budget support programmes have predominantly a positive or mixed effect on the topics laid out in the CEF.

Table 8 presents the coverage of budget support effects in the evidence. It displays the number of reports that mention budget support effects per topic. The yellow cells highlight those topics for which fewer than 10 reports provide evidence. (This is not to be confused with an absent or negative effect of budget support on these topics.) The blue cells mark all topics covered by 10 or more reports.

As can be seen from the first two columns in Table 8, all but two topics have sufficient evidence. Thus, on the generic level, most effects of budget support programmes are sufficiently covered by evidence. The exceptions are effects on anti-corruption agencies and corruption legislations. However, as illustrated by the last four columns, the evidence attributes most observed effects to budget support programmes on a general level, but a specific attribution of effects to individual budget support inputs is rarely covered by the evidence.

^{xlvii} The number of sources indicated in Table 11, and subsequent tables reporting the results of the quantitative analysis, can differ from the number of codings per effect direction analysed in the qualitative analysis. While the quantitative analysis is based on the total of all coded effects, in the qualitative analysis a) only one coding on the same effect per source is included, b) only codings with a plausible attribution to budget support are included, c) where possible, weak and strong positive effects are distinguished. As a consequence, the number of sources covering positive and negative effects of budget support is higher in the quantitative than in the qualitative analysis. To ensure the reliability of evidence, findings on budget support effects were based on qualitative findings, while the quantitative analysis was used for the description of coverage.

Table 8: Coverage of evidence for all budget support effects^{xlviii}

		Total	Overall BS Programmes	Financial Inputs	Non-Financial Inputs		
					Policy Dialogue	Conditionality	TA / CD
Aid-effectiveness Principles	Alignment	38	37	(8)	(9)	(6)	(4)
	Harmonization	44	41	(4)	14	17	(5)
	Ownership	46	36	(6)	16	23	(2)
	Predictability	47	29	30	(2)	13	(1)
	Transaction Costs	38	34	14	(6)	(4)	(0)
Government Expenditure	Level of Public Spending	38	20	30	(2)	(3)	(0)
	Composition of Public Spending	28	19	18	(4)	(2)	(0)
	Domestic Revenue Mobilization	34	24	12	(2)	(7)	(3)
Public Financial Management	Public Financial Management	58	52	(9)	19	15	29
	Budget Management	34	27	(7)	(7)	(7)	(5)
	Budget Transparency	29	24	(3)	(3)	(4)	(4)
Corruption	Anti-corruption Agency	(9)	(9)	(2)	(1)	(1)	(2)
	Corruption Legislation	(5)	(5)	(0)	(0)	(0)	(0)
	Corruption Level	19	15	(3)	(1)	(2)	(1)
	Corruption Policies	13	10	(0)	(2)	(3)	(1)
Domestic Accountability	Domestic Accountability	44	42	(8)	(9)	(7)	(8)
	Supreme Audit Institutions	17	15	(6)	(2)	(3)	(4)
	Role of Parliament	21	20	(4)	(4)	(1)	(2)
	Civil Society Organizations	25	24	(4)	(8)	(1)	(2)
	Governance	31	27	(1)	(8)	(6)	(4)
Economic Performance and Income Poverty	Macroeconomic Performance	31	28	(8)	(3)	(4)	(1)
	Income Poverty	40	36	17	(7)	(4)	(1)
Service Delivery and Non-income Poverty	Public Service Delivery	47	40	28	10	(7)	10
	Administration for Service Delivery	25	18	(6)	(5)	(5)	(6)
	Non-income Poverty	33	32	(9)	(2)	(0)	(0)

Table 9 shows the effect direction for all effects that are sufficiently covered by evidence.

The statements included in this table are made in more than 10 sources and are therefore considered as sufficiently covered.

The interpretation of the “traffic light” is as follows: *red* means there is no budget support effect on the output, outcome or impact, *yellow* means there is a mixed effect and *green* means there is a positive effect. *White* means that the attribution of the observed effects to budget support is not possible.

^{xlviii} In Table 8, the coverage of budget support effects on outputs, outcomes and impacts is categorized as follows: coverage in fewer than 10 reports (number of reports in brackets) yellow; 10 – 34 reports (sufficient coverage) light blue; 35 – 49 reports (strong coverage) blue; 50 – 95 reports (very strong coverage) dark blue.

Table 9: Evidence on the effectiveness of budget support

	Theme	Finding	Effectiveness
Aid-effectiveness Principles	Alignment	Budget support programmes increase alignment of aid with national policies, although it is still described as weak by the studied evaluations and reports.	●
	Harmonization	Budget support, in particular non-financial inputs, increases harmonization among donors when there is a set of common goals and priorities established through coordination process.	●
	Ownership	Budget support fosters the ownership of the partner governments, in particular when there is a positive effect of policy dialogue on ownership. However, conditionality has mixed effects.	●
	Predictability	Budget support does not improve overall predictability due to late approval and disbursement of tranches in response to the (non-)adherence to conditionalities.	●
	Transaction Costs	Transaction costs increased initially, but decreased as processes became routinely established. This effect diminished once budget support suspensions started to be enforced.	●
Government Expenditure	Level and Composition of Public Spending	Budget support funding increases the level of public spending, specifically in social sectors such as health and education.	●
	Domestic Revenue Mobilization	Access to budget support funds does not reduce the mobilization of domestic revenues, thus there is no evidence for a “crowding out effect”. However, the evidence points to insignificant improvements related to DRM.	●
Public Financial Management	Public Financial Management	Budget support programmes have a positive effect on improving public financial management of recipient countries, in particular on budget formulation and planning, auditing and budgeting, and comprehensiveness and transparency of the budget.	●
	Budget Management		●
	Budget Transparency		●
Corruption	Corruption Level	Overall, budget support has no negative effect on the corruption level, and has a positive effect on strengthening the mechanisms for monitoring corruption. However, prosecution of corruption cases is still lagging and there are very few convictions.	●
	Corruption Policies	Budget support does not have a negative effect on corruption policies. The evidence shows mixed effects (no or positive effects) on corruption policies.	●
Domestic Accountability	Supply Side	Budget support programmes improve domestic accountability, in particular the supply side (“duty-bearers”). It substantially augments the role of SAI.	●
	Demand Side	There are mixed effects of budget support on the role of Parliament and civil society organizations (“rights holders”): there is no proof of either a positive or a negative effect of budget support.	●
	Governance	Budget support has a positive effect on intra-governmental accountability structures, but has mixed effects on decentralization.	●
Economic Performance and Income Poverty	Macroeconomic Performance	Budget support reinforces pre-existing macroeconomic stability and enables repayment of domestic debt to a certain extent, but direct attribution to the influence of budget support beyond this is unclear.	●
	Income Poverty	Evidence on the effects of budget support on income poverty is mixed: most of the sources find no or only slightly positive effects and the attribution to budget support is difficult.	○
Service Delivery and Non-income Poverty	Public Service Delivery and Administration	The quantity of provided services increases due to budget support programmes, however the quality does not improve.	●
	Non-income Poverty	There are positive effects on non-income poverty in recipient countries, but effects are not attributed to budget support only.	○

Table 10: Codes on aid-effectiveness principles

Topic	Code	Definition
Aid-effectiveness Principles	Alignment	Coordination and consistency of aid with government priorities
	Harmonization	Harmonization of donor's aid policies and procedures
	Ownership	Recipient governments formulating and implementing own priorities and strategies for national development and poverty reduction
	Predictability	Certainty about amounts and timing of aid disbursements
	Transaction costs	Transaction costs per Euro of aid disbursed, including costs for preparation, negotiation, disbursement and monitoring

4.1 Aid-effectiveness principles

The aid-effectiveness principles as defined in this evaluation synthesis (see 2.1 for an explanation) include *alignment*, *harmonization and ownership* as well as *predictability* and *reduction of transaction costs*.^{xlix} The specific definitions of the principles as used in this evaluation synthesis are stated in Table 10.

Summary finding: Positive effects on alignment, harmonization, and ownership, but low coverage for specific budget support inputs

The effects of budget support programmes on aid-effectiveness principles are sufficiently covered by evidence on a generic level, with the highest coverage for predictability and the lowest coverage for alignment. Conclusions on these effects are presented in the paragraph below. The specific attribution of changes in aid-effectiveness principles to budget support inputs is insufficiently covered by evidence for most of the codes. The attribution of changes in alignment to financial and non-financial budget support effects and of changes in transaction costs to non-financial effects is lacking. For all five aid-effectiveness principles, the effect of the input TA/CD is insufficiently covered.

Budget support programmes have a positive effect on alignment, harmonization and ownership, although budget support's effectiveness depends on the number and quality

of donors' common interests and the adherence of partner governments to the conditionalities attached to budget support programmes. Financial inputs of budget support do not have a positive effect on predictability. This is due to the delay of aid disbursements and uncoordinated decisions of donors following breaches of the conditionalities by the partner governments. The non-financial inputs policy dialogue and conditionality have a positive effect on harmonization, but do not contribute to increased ownership. Conditionality might even have a negative effect on predictability. Transaction costs increased after introducing budget support, but declined when joint processes were established. When donors withdraw from budget support, the suspension of tranches increases transaction costs again, as only small sums are paid out in relation to the operating expenses.

4.1.1 Expected effects

According to the generic intervention logic, the effectiveness of budget support depends on how the programme is implemented, in particular the extent to which donors adhere to aid-effectiveness principles when implementing budget support. Budget support is expected to ensure alignment and foster ownership by adhering to partner countries' own national development strategies and by using the partner governments' systems and procedures. With donors jointly providing non-financial inputs and coordinating disbursements and monitoring activities, budget support is expected to foster harmonization of donor procedures. By providing the funding

^{xlix} Note that these principles are derived from the CEF and are not the same as those formulated in the Paris Declaration on Aid Effectiveness.

Table 11: Budget support effects on alignment

Inputs		Total	BS Programme	Funds	Policy dialogue	Conditionality	TA / CD
Alignment	# Sources	38	37	8	9	6	4
	Total effect codings	91	61	8	11	6	5
	Positive effect	76	48	7	11	6	4
	Absent effect	6	5	0	0	0	1
	Negative effect	1	1	0	0	0	0
	Improving effect	5	4	1	0	0	0
	Deteriorating effect	3	3	0	0	0	0

through government channels and by linking the disbursements to progress against an annual performance assessment, budget support is expected to improve predictability. Moreover, a reduction of transaction costs is expected, as budget support is provided in a way that does not require separate project management, planning or monitoring systems, but uses existing partner country systems.⁵⁸

4.1.2 Alignment

Coverage

In the sample, 38 sources cover the total effect of budget support on alignment (see Table 11). Compared to the other aid-effectiveness principles, alignment has the lowest coverage (together with transaction costs). Generic effects of budget support programmes on alignment are reported by 37 sources. The effects of each of the specific inputs are covered in fewer than 10 sources: the effects of funding on alignment are covered in eight sources, of policy dialogue in nine sources, of conditionality in six sources and of TA/CD in four sources.

That said, too few sources report on the input-specific attribution of changes in alignment to budget support funding. Whether or not the effects hypothesized in the CEF apply in reality, and specifically to the promotion of government ownership and accountability through aligned budget support funding, cannot be answered. Also, the reports do not focus on the input-specific attribution of changes in alignment to non-financial inputs. Due to this evidence gap, it is not known if

and how budget support improves coordination and consistency with government priorities. Without knowledge of the specific impact of budget support funding and non-financial inputs on alignment, no lessons can be learned on how to increase the effectiveness of this or similar aid modalities.

Established evidence

In line with the expectation formulated in the intervention logic, budget support increases the alignment of aid with the national policies of partner countries. This effect is covered in 29 sources⁵⁹ and found in 16 countries¹ that received budget support.⁶⁰ The positive effect of budget support refers to the progress in aligning aid to national development plans and poverty reduction strategies,ⁱⁱ ⁶¹ in increasing the use of government systems to implement the aid modality⁶² and in aligning the aid to the specific economic and institutional context.⁶³

The evidence from these sources suggests that although donors' alignment with government priorities improved as a result of budget support, the overall alignment of budget support programmes was insufficient. Reasons mentioned in the sources are the weak alignment of donors' aid policies with the priorities and national government strategies of the partner countries at sub-national level, and insufficiencies in the integration of existing accountability systems and supreme audit institutions (SAI). Another reason mentioned for shortcomings in alignment is that the design and

¹ Burkina Faso, Dominican Republic, Ethiopia, Ghana, Malawi, Mali, Mozambique, Nicaragua, Rwanda, Santa Lucia, South Africa, Tanzania, Tunisia, Uganda, Vietnam, and Zambia.

ⁱⁱ E.g. in Burkina Faso, Dominican Republic, Ghana, Malawi, Mali, Mozambique, Nicaragua, Rwanda, Santa Lucia, Uganda, Vietnam, and Zambia.

Table 12: Budget support effects on harmonization

Inputs		Total	BS Programme	Funds	Policy dialogue	Conditionality	TA / CD
Harmonization	# Sources	44	41	4	14	17	5
	Total effect codings	123	73	4	20	19	7
	Positive effect	84	55	1	15	10	3
	Absent effect	19	6	1	3	5	4
	Negative effect	17	10	2	1	4	0
	Improving effect	1	1	0	0	0	0
	Deteriorating effect	2	1	0	1	0	0

implementation of the budget support programmes is still strongly donor driven.⁶⁴

4.1.3. Harmonization

Coverage

Effects on harmonization are strongly covered in 44 sources with 123 codings (see Table 12). The generic attribution of changes in harmonization to budget support programmes is covered in 41 sources. The input-specific attribution of changes in harmonization to the individual budget support inputs policy dialogue (14 sources) and conditionality (17 sources) is also covered in a sufficient number of sources. The effects of funds are only covered in four sources and of TA/CD in five sources. The effects of these inputs on harmonization are not deemed to be sufficiently covered by empirical evidence to derive generalizable conclusionsⁱⁱⁱ on their effectiveness. Evidence on the particular effects of budget support funds and TA/CD on harmonization is thus not available, and the expectation that budget support funding increases the harmonization of joint donor procedures⁶⁵ in, for example, monitoring and disbursement, cannot be empirically corroborated.

Established evidence

Budget support improves harmonization among donors, but only under certain conditions. Progress in donor harmonization is reported in 19 sources⁶⁶ covering 13 countries.ⁱⁱⁱ Specifically, according to the evidence, budget support improves the

harmonization of the donors' aid policies and procedures (covered in 13 sources).⁶⁷ Yet, diverging priorities of donors, such as different priority sectors and a lack of common goals, undermine donors' willingness to provide budget support disbursements in a harmonized way, and weaken the effectiveness of budget support on harmonization.^{iv 68}

Harmonization is further undermined by the uncoordinated disbursement decisions of individual donors, an issue that is consistently reported in the reviewed body of evidence.⁶⁹ This is particularly evident for country cases^{iv} where financial inputs are disbursed following additional bilateral agreements, and breaches of these agreements lead to unilateral decisions by individual donors.⁷⁰ Thus, progress in harmonization is evident in the overall procedures for budget support disbursements, but less so regarding individual disbursement decisions and practices.

With respect to the non-financial inputs policy dialogue and conditionality, the evidence on effects on harmonization is mixed.⁷¹ Overall, policy dialogue is found to be effective for harmonization, even in cases of large and heterogeneous donor groups, as it provides an institutionalized setting for donor coordination and the harmonization of policies and practices.^{iv 72} However, the continuation of bilateral meetings in addition to the policy dialogue, for example individual missions to discuss implementation specifics and performance assessment, undermine the positive effect of policy dialogue

ⁱⁱⁱ The term "generalizable conclusions" implies results that are backed by strong empirical evidence; it is not meant to describe universally accepted results.

ⁱⁱⁱ Burkina Faso, Ghana, Malawi, Mozambique, Nicaragua, Rwanda, Sierra Leone, Tanzania, Tunisia, Uganda, Vietnam, and Zambia.

^{iv} Such as in Afghanistan, Burundi, CAR, DRC, Ethiopia, Nepal, Sierra Leone, Tunisia, Uganda, and Zambia.

^{iv} E.g. Malawi, Nicaragua, Ruanda, Uganda, and Zambia.

^{iv} Such as in Ghana, Malawi, Mali, Mozambique, Rwanda, Tunisia, Uganda, and Zambia.

Table 13: Budget support effects on ownership

Inputs		Total	BS Programme	Funds	Policy dialogue	Conditionality	TA / CD
Ownership	# Sources	46	36	6	16	23	2
	Total effect codings	133	76	6	17	32	2
	Positive effect	91	52	5	15	18	1
	Absent effect	24	15	0	1	7	1
	Negative effect	12	5	1	1	5	0
	Improving effect	5	3	0	0	2	0
	Deteriorating effect	1	1	0	0	0	0

and conditionality on harmonization.⁷³ Conditionality, predominantly the PAF, also fosters harmonization as it establishes a joint approach to a government performance assessment and defines common indicators for measuring this performance.⁷⁴ However, insufficient harmonization in indicators of the PAF and donors' quest for different priorities are limiting the effects of budget support on harmonization.^{lvii 75}

4.1.4 Ownership

Coverage

Budget support effects on ownership are covered by 46 sources and found in 133 codings (see Table 13), the largest number of codings among the aid-effectiveness principles. The generic effects of budget support programmes, unspecified for inputs, on ownership are reported in 36 sources. The effect of policy dialogue is covered in 16 sources and that of conditionality in 23 sources. Funding is discussed in only six sources and TA/CD in two sources. Therefore, no generalizable conclusions on the effect of funding and TA/CD on ownership can be drawn and it remains unclear if TA/CD is indeed effective in giving the recipient countries a greater say in the design, policy dialogue, and conditionalities⁷⁶ of budget support programmes. No lessons can thus be drawn to inform future programmes of budget support or related aid modalities with respect to the impact of funding and TA/CD on ownership.

Established evidence

Budget support fosters ownership by the partner governments in terms of formulating and implementing their own priorities and strategies for national development and poverty reduction. This positive effect is found in 18 sources⁷⁷ covering 15 country cases.^{lviii} In four countries, no effects of budget support on ownership are found.^{lix 78} The evidence shows that the effectiveness of budget support depends, among other factors, on the degree of ownership the partner governments assume towards development plans and poverty-reduction strategies.⁷⁹ Also, the degree of ownership seems to be connected to the content of reform and the politics of the government in power.⁸⁰ While positive effects were found mostly in the early period of budget support programmes, the evidence suggests that the effect on ownership diminished over time, coinciding with a higher frequency of budget support suspensions. Some examples are Rwanda, Tanzania, Uganda, Vietnam, and Zambia.⁸¹

While there is sufficient coverage of policy dialogue and conditionality as non-financial inputs of budget support in the sources,⁸² the reported direction of effects is inconsistent across these sources. The effectiveness of conditionality is positively influenced when the PAF is derived from the national development plans of the recipient country, and when the PAF process is strongly led by the government.⁸³ The evidence suggests that, among the PAF indicators, alignment with planning and budgeting strategies is particularly effective in increasing ownership.⁸⁴ Other sources, however, question

^{lvii} Malawi, Rwanda, Uganda, and Zambia.

^{lviii} Benin, Burkina Faso, Dominican Republic, Ghana, Malawi, Mali, Mozambique, Nicaragua, Rwanda, South Africa, Tanzania, Uganda, Vietnam and Zambia.

^{lix} In the cases of Ethiopia, India/Andhra Pradesh, Palestine, and Santa Lucia.

Table 14: Budget support effects on predictability

Inputs		Total	BS Programme	Funds	Policy dialogue	Conditionality	TA / CD
Predictability	# Sources	47	29	30	2	13	1
	Total effect codings	104	36	43	3	21	1
	Positive effect	60	22	31	1	6	0
	Absent effect	13	8	2	0	3	0
	Negative effect	25	4	7	1	12	1
	Improving effect	4	0	3	1	0	0
	Deteriorating effect	2	2	0	0	0	0

how effective and sustainable the observed increases in ownership are.⁸⁵

The positive effects of policy dialogue in fostering ownership for development policies are undermined by a lack of collaboration between donors and partners,⁸⁶ and by the exclusive and government-focused nature of the dialogue. As the participation of wider civil society or even parliamentarians in the dialogue is not always guaranteed, budget support does little to induce a shift from narrow government ownership of national development plans and policies to wider democratic ownership for these processes.⁸⁷ ^{lx}

4.1.5 Predictability

Coverage

The effect of budget support on predictability^{lxi} is covered in 47 sources, the highest coverage being among the aid-effectiveness principles (see Table 14). Regarding the generic attribution of changes in predictability to budget support programmes, the effectiveness of predictability is covered in 29 sources. Regarding the specific inputs, the predictability of funding is covered in 30 sources, effects of conditionality are reported in 13 sources, effects of policy dialogue in two sources and effects of TA/CD in one source. As there is insufficient coverage on policy dialogue and TA/CD, no generalizable conclusions on these effects can be derived regarding the effectiveness of these two inputs.

Established evidence

Although a considerable number of sources cover effects on predictability (47 sources), the direction of the effects found is inconsistent across sources. There is evidence for positive effects of budget support on predictability, but these positive effects are mostly described as weak (12 sources).⁸⁸ Other sources report a development over time, with no changes in predictability at the beginning of programmes, but increases in predictability at later stages of the budget support programme cycle (12 sources).⁸⁹ No improvements or even negative effects are reported in 16 sources.⁹⁰ The inconsistency in effect direction indicates that the effectiveness of budget support on predictability is highly dependent on context conditions.

Weak positive effects of budget support on the predictability of aid are reported for seven countries.^{lxii} ⁹¹ In most of the sources, such improvements in predictability are defined as a slight decrease in deviations of budget support disbursements from commitments. Conditions found for positive effects of budget support programmes are the disbursement of budget support funds following a multi-year arrangement (three-year or five-year basis), better and early notification of disbursements by donors, and effective mutual accountability to hold donors responsible for their commitments.⁹²

Regarding changes over time, improvements in aid

^{lx} E.g. Burkina Faso, India/Andhra Pradesh, and Mozambique.

^{lxi} Although the reports apply different comparisons, most reports compare the predictability of budget support disbursements with project aid and in fewer cases, with basket funding. For a conservative assessment, this evaluation describes improvements in predictability in comparison to project aid.

^{lxii} Benin, Burkina Faso, Ethiopia, Ghana, Mali, Senegal, and South Africa.

Table 15: Budget support effects on transaction costs

Inputs		Total	BS Programme	Funds	Policy dialogue	Conditionality	TA / CD
Transaction Costs	# Sources	38	34	14	6	4	0
	Total effect codings	90	55	20	10	5	0
	Positive effect	51	33	14	2	2	0
	No effect	12	6	2	3	1	0
	Negative effect	22	13	4	3	2	0
	Improving effect	3	2	0	1	0	0
	Deteriorating effect	2	1	0	1	0	0

predictability seem to develop with the progress in budget support programmes.^{lxiii} 93 However, problems in fulfilling disbursement conditions, which lead to delayed or suspended instalments, undermine predictability. This effect is reinforced by diverging donor responses to breaches of underlying principles by the partner government, and resulting individual decisions about budget support suspensions.⁹⁴

Also, the evidence finds that unrealistic and inadequately communicated conditionalities result in unpredictable disbursements.⁹⁵ Such negative effects of conditionality on predictability are covered in 10 sources⁹⁶ and reported in the case of nine countries.^{lxiv}

No overall improvements on predictability are found in the case of 11 countries.^{lxv} Late confirmations of tranche releases on the part of donors, discrepancies among donors in the timing of disbursements, and late timing of budget support disbursements within the fiscal year lead to the ineffectiveness of budget support in promoting predictability. The evidence shows that late disbursements have negative, unintended effects on the policy process, leading to unplanned domestic borrowing and to financial challenges for sector ministries.⁹⁷ Moreover, frequent changes in disbursement conditions and the use of performance-based tranches in budget support programmes further reduce budget supports' ability to positively influence predictability.⁹⁸ The approval of tranche releases on an annual basis does not promote multi-year

planning and eventually the early incorporation of funds in the budget process.⁹⁹

4.1.6 Transaction costs

Coverage

The total number of effects of budget support on transaction costs^{lxvi} is covered in 38 sources (see Table 15). The generic attribution of changes in transaction costs to budget support programmes is examined in 34 sources. Regarding the individual inputs, the effects of funding on transaction costs are covered in 14 sources. For the non-financial inputs, the coverage is fewer than ten reports, as effects of policy dialogue are reported in six sources, of conditionality in four, and of TA/CD in no sources. Therefore, no lessons can be derived from the implementation of non-financial inputs on the reduction of transaction costs, but generalizable conclusions can be drawn for the effect of financial inputs and the GBS programmes on transaction costs.

Established evidence

The evidence supports a positive effect on the reduction of transaction costs, but effects fall short of expectations (22 sources).¹⁰⁰ The evidence only covers effects on transaction costs on the side of the partner government. Effects on the side of the donors are insufficiently covered. This section thus only describes the side of the partner government.

The evidence shows time-varying effects of budget support on

^{lxiii} E.g. in the cases of Mozambique, Nicaragua, and to some extent Rwanda.

^{lxiv} Afghanistan, Malawi, Mozambique, Nicaragua, Palestine, Rwanda, Sierra Leone, Tanzania, and Uganda.

^{lxv} Afghanistan, Burundi, CAR, DRC, India/Andhra Pradesh, Malawi, Nepal, Sierra Leone, Uganda, Vietnam, and Zambia.

^{lxvi} The reports use slightly different definitions of transaction costs, but most include the costs for preparation, negotiation, disbursement and monitoring of transactions. This definition was applied in this evaluation synthesis.

Table 16: Government expenditure codes

Topic	Code	Definition
Government Expenditure	Level of public spending	Total amount of government spending
	Composition of public spending	Allocation of government spending to different sectors
	Domestic revenue mobilization	Acquiring additional sources for public/state income

transactions costs for partner governments, with initial increases in the transaction-cost burden, and subsequent reductions as processes become routinely established. This is, for example, the case in Ghana, Tanzania, Uganda, Mozambique, and Nicaragua.¹⁰¹

Despite some reduction in transaction costs, the evidence reports on different reasons why the decrease in transaction costs does not meet expectations.¹⁰² Factors hindering a stronger reduction are found to be the substantial capacity needs of both partners and donors for an institutionalized dialogue, accurate monitoring and reporting. In contrast to the assumption that budget support lessens the need for separate project management and monitoring,¹⁰³ the evidence shows that capacity needs are reinforced by additional assessments and missions of individual agencies.¹⁰⁴ Also, a further reduction in transaction costs is limited by the amounts disbursed and the operation mode for disbursements. The increased use of variable tranches by donors, the low share of budget support of total aid volumes towards the end of the budget support programmes, and the necessity of managing short-term volatilities in disbursements are such factors.¹⁰⁵

As a result of financial inputs of budget support on transaction costs, the evidence suggests that in certain phases of budget support disbursements, the total volume of aid disbursed is higher and transaction costs per Euro of aid disbursed are found to be lower than for project aid.¹⁰⁶ However, when the amount of budget support funding decreased and suspensions started to be used by donors more frequently, this positive effect on transaction costs diminished.¹⁰⁷

4.2 Government expenditure

The topic government expenditure is sub-divided into the specific outcomes level and composition of public spending and domestic revenue mobilization (DRM).¹⁰⁸ The definitions of these codes are stated in Table 16.

Summary finding: Budget support (funding) increases public spending in social sectors – deficit in describing effects of non-financial inputs

The attribution of changes in government expenditure to budget support programmes on a generic level is sufficiently covered by evidence, with the highest coverage for the level of public spending and the lowest coverage for the composition of public spending. The effects of budget support funding are sufficiently covered for all codes on government expenditure, while not a single non-financial input and its effect on government expenditure is sufficiently covered. Due to this substantial deficit, no generalizable conclusions on the expected positive influence of non-financial inputs going beyond the sole financing function of budget support can be drawn.

Budget support, especially the financial input, increases public spending. This increase is mostly apparent in social sectors such as health and education. Flexible funding in combination with performance assessments seems to positively influence the composition of public spending. There is no evidence that access to budget support funds reduces the mobilization of domestic revenues. At the same time, there are no, or only slight, improvements in DRM, putting at risk the sustainability of budget support effects on (social) public spending.

¹⁰¹ E.g. in Ghana, Nicaragua, Tanzania, Uganda, Vietnam, and Zambia.

¹⁰⁸ The income side is covered in this section – and not in the section PFM – as a counterpart to government expenditure and in order to discuss the crowding-out effect.

Table 17: Budget support effects on level of public spending

Inputs		Total	BS Programme	Funds	Policy dialogue	Conditionality	TA / CD
Level of Public Spending	# Sources	38	20	30	4	2	0
	Total effect codings	97	30	61	4	2	0
	Positive effect	81	25	52	3	1	0
	No effect	5	1	4	0	0	0
	Negative effect	4	1	2	0	1	0
	Improving effect	1	0	1	0	0	0
	Deteriorating effect	6	3	2	1	0	0

4.2.1 Expected effects

Budget support programmes, in particular their financing function, are expected to have substantial effects on the level and composition of public spending.¹⁰⁸ According to the generic intervention logic, the provision of resources through the partner government's own budget facilitates the government in planning the allocation of funds systematically and in aligning allocation with its strategic priorities. Budget support as an aid modality is also expected to increase a government's spending capacity.¹⁰⁹ Another expected effect of budget support programmes is public spending more strongly orientated towards the poor through the promotion of fiscal and macroeconomic stability, including the aggregate allocation of resources, higher quantities of aid "on-budget", higher allocation to priority sectors and higher consistency between recurrent and investment sides of the budget.¹¹⁰ A purported risk of budget support is that the provision of on-budget funding may undermine incentives for the partner governments to mobilize sufficient domestic revenue to finance national policies, and thus crowd out domestic revenue instead of providing additional resources for public spending.¹¹¹

4.2.2 Level of public spending

Coverage

The level of public spending has the best coverage among the four outcomes under the topic government expenditure. It is discussed in 38 sources, with 97 coded text sections (see Table 17). A generic attribution of changes in the level of public spending to budget support programmes is described in 20

sources. The sources refer mostly to the systemic effect of budget support funds, with 30 sources mentioning the effects of funding on government expenditure. Non-financial inputs are rarely mentioned in the sources, with only four sources mentioning effects of policy dialogue, two sources mentioning conditionality, and none of the sources mentioning TA/CD. Since the effects of non-financial inputs on the level of public spending are covered in fewer than 10 sources, no generalizable conclusions on the direct effect of non-financial inputs can be drawn, whereas the sufficient coverage of the systemic effect^{lxix} of budget support funding allows for conclusions on the effectiveness of budget support in this regard.

Established evidence

Overall, budget support is found to increase the level of public spending in recipient countries. Such an increase in public spending is reported in 14 sources, covering 12 countries.^{lxx lxxii} When budget support programmes were suspended, in some country cases^{lxxi} the gained fiscal space was completely eliminated.¹¹³ The positive effects are also found to depend on the degree of predictability and volatility of budget support funds,¹¹⁴ donor pressure to increase social spending, and the scale and flexibility of funding in relation to policy goals.¹¹⁵ The evidence suggests that a combination of high flexibility in financing and performance assessments for national strategies increases the effectiveness of public spending for service delivery, for example, as constraints in the allocation of off-budget aid could be compensated and the budget more flexibly managed.¹¹⁶

^{lxix} A systemic effect is generated by a flow of funds through the government's own system (Nilsson, 2004).

^{lxx} Burkina Faso, Burundi, Ghana, Mali, Morocco, Mozambique, Rwanda, Sierra Leone, South Africa, Tanzania, Uganda, and Vietnam

^{lxxi} Burundi, Malawi, Rwanda, and Uganda.

Table 18: Budget support effects on composition of public spending

Inputs		Total	BS Programme	Funds	Policy dialogue	Conditionality	TA / CD
Composition Public Spending	# Sources	28	19	18	2	3	0
	Total effect codings	63	28	29	2	4	0
	Positive effect	51	20	27	2	2	0
	No effect	5	5	0	0	0	0
	Negative effect	6	3	1	0	2	0
	Improving effect	0	0	0	0	0	0
	Deteriorating effect	1	0	1	0	0	0

It is the funding provided by the budget support programmes that leads to the provision of additional funds reflected in the national budget. The specific results of the funding effect presented in the evidence are an increase in fiscal space and more funding for the government to finance priority sectors.¹¹⁷ As more funding appears to translate into more spending, this finding suggests that no crowding out of domestic revenues takes place (see Box 3).

Box 3 No evidence of “crowding out” of domestic revenue by budget support

One of the suspected risks of budget support provision was that this aid modality would reduce incentives for governments to raise domestic revenues and, as a consequence, crowd out domestic revenue: the “crowding out effect”.¹²⁹ However, the sources analysed in this evaluation synthesis do not offer evidence that access to budget support funds reduced the mobilization of domestic revenues. In particular, the absence of such crowding out effects was reported in 12 countries: Burkina Faso, Ghana, Malawi, Mali, Mozambique, Nicaragua, Rwanda, Sierra Leone, Tanzania, Uganda, Vietnam, and Zambia.¹³⁰

4.2.3 Composition of public spending

Coverage

The composition of public spending as an outcome of budget support is covered in 28 sources and 63 text sections, which is the weakest coverage found among the different codes on government expenditure. There is sufficient coverage on the generic effect of budget support programmes (19 sources) and funding as a specific budget support input (18 sources) on the composition of public spending. There is little to no evidence on the three non-financial inputs, with two sources mentioning effects of policy dialogue and three sources mentioning effects of conditionality (see Table 18). Somewhat surprisingly, the analysis thus finds that what, in some respects, can be considered the centrepiece of the intervention logic of budget support – namely that the combination of financial resources with non-financial inputs aimed at improving the use of those resources – is barely covered in the body of evidence. The assumption that the combination of policy dialogue, conditionality and TA/CD has a positive influence on the spending policy, and thus supports pro-poor spending beyond the sole financing function of budget support, cannot be confirmed or disconfirmed.

Established evidence

A strong body of evidence indicates a positive link between budget support and the composition of public spending (48 text sections in 27 sources).¹¹⁸ Evidence for higher social-sector spending is found in nine countries.^{1xxii 119} The evidence finds that additional funds provided through budget support are important for the implementation of national development

^{1xxii} Burkina Faso, Burundi, Ghana, Malawi, Mozambique, Nicaragua, Rwanda, Uganda, and Vietnam.

Table 19: Increased spending in pro-poor sectors

Sector	Country	Source
Health	Burkina Faso, Burundi, Ghana, Ethiopia, Malawi, Mali, Mozambique, Nicaragua, Niger, Rwanda, Sierra Leone, Tanzania, Uganda, Vietnam, Zambia	ADE, 2015; AfDB, 2011b; Alavuotunki, 2015; de Kemp et al., 2011a; Lawson, 2007; Lawson et al., 2007; Lawson et al., 2013; Lawson et al., 2014; World Bank, 2008; World Bank, 2015a; World Bank, 2015b
Education	Burundi, Ethiopia, Ghana, Mali, Morocco, Mozambique, Nicaragua, Rwanda, Sierra Leone, Tanzania, Uganda, Vietnam, Zambia	ADE, 2015; AfDB, 2011b; de Kemp et al., 2011a; Dijkstra et al., 2012; Lawson, 2007; Lawson et al., 2006; Lawson et al., 2007; Lawson et al., 2013; Lawson et al., 2014; ODI, 2009; Purcell et al., 2006; Thunnissen/Morillon, 2014; World Bank, 2008; World Bank, 2015a; World Bank, 2015b
Agriculture	Burundi, Tanzania	ADE, 2015; AfDB, 2011b; Lawson et al., 2013
Infrastructure	Burundi, Tanzania, Uganda	ADE, 2015; AfDB, 2011b; Bogetic et al., 2015; Lawson et al., 2013
Water and Sanitation	Tanzania, Uganda	AfDB, 2011b; Lawson et al., 2013; World Bank, 2015a

plans.¹²⁰ Some country cases state that the access to new discretionary funding is used to increase spending in social sectors rather than to reallocate funds to other sectors.¹²¹ In particular, the evidence suggests that certain expenditures, such as (non-salary) recurrent expenditure and expenditures for intergovernmental transfers, would thus have been difficult to finance without the additional funds provided by budget support programmes. Budget support contributes to reducing fiscal deficits and thus macroeconomic stability by allowing additional development-oriented spending within the limits of an expanded resource envelope.

The increases in spending are found to be predominantly in health and education (see Table 19). Increased expenditure in the health system was found in 15 country cases, 11 of them being low-income countries (LICs). An increase in spending on education is reported in 13 country cases, eight of them being LICs. In contrast, allocation of increased funds into agriculture, infrastructure (e.g. roads) and water and sanitation is only found in two to three country cases, and the evidence base is fewer than 10 sources.¹²² Increases in social sector spending seem to be limited to the health and education sectors. Although most of the sources stating this effect report increases in pro-poor spending, not all of them document if a comprehensive poverty-spending analysis has been conducted. Thus, it is not clear if the observed increases in social sector spending are equivalent to an increase in pro-poor spending.

Overall, the sources draw a distinctly positive picture with regard to budget support's influence on social spending and in most cases do not report on possible negative effects, which translates into the positive picture drawn in Table 19. The very few negative effect codings refer to the negative influence of budget support suspensions on the fiscal situation as the suspensions prompted a change in governments' investment priorities towards the productive sector, with a subsequent decline in allocations to social sectors.¹²²

4.2.4 Domestic revenue mobilization

Coverage

Effects of budget support on DRM are covered in 34 sources and 70 text sections (see Table 20). Generic effects of the budget support programmes are covered in 24 sources, and 12 sources mention the effects of funding. The effects of non-financial inputs on DRM are not sufficiently covered (policy dialogue covered in two sources, conditionality in seven sources and TA/CD in three sources), thus no generalizable lessons can be derived on the effects of non-financial inputs on DRM.

Established evidence

Most of the reports indicate no effect of budget support on DRM (32 text sections in 18 sources).¹²³ Although there is sufficient evidence that the provision of budget support does not diminish DRM (see Box 3), in only a few country cases¹²⁴ do the sources identify positive effects of budget support

¹²⁰ The sectors listed in Table 19 are the sectors being discussed in more than one source regarding the composition of public spending.

¹²¹ Burkina Faso, Burundi, Mozambique, Tanzania, and Tunisia.

Table 20: Budget support effects on domestic revenue mobilization

Inputs		Total	BS Programme	Funds	Policy dialogue	Conditionality	TA / CD
Domestic Revenue Mobilization	# Sources	34	24	12	2	7	3
	Total effect codings	70	43	14	3	7	3
	Positive effect	23	16	3	1	2	1
	No effect	39	24	8	2	3	2
	Negative effect	8	3	3	0	2	0
	Improving effect	0	0	0	0	0	0
	Deteriorating effect	0	0	0	0	0	0

programmes on revenue mobilization.¹²⁴ A poor record on DRM is reported in several countries (Burkina Faso, Sierra Leone, Tanzania, and Uganda), such as low tax and non-tax revenues.¹²⁵

However, the negative dynamic is not attributed to budget support programmes, with the exception of Sierra Leone. In this case the evidence suggests that budget support might have provided disincentives to increase DRM.¹²⁶

The persistent low-revenue mobilization in many countries undermines the sustainability of budget support effects. When budget support disbursements started to decrease, the remaining available funds were insufficient to keep the increased spending in social sectors up. As a consequence, domestic borrowing and redirection of funds away from non-poor sectors started to increase, such as occurred in Malawi and Uganda.¹²⁷ Low levels of domestic revenue thus enhanced the negative effect of budget support suspensions on the level and composition of public spending.¹²⁸

Table 21: Public financial management codes

Topic	Code	Definition
Public Financial Management (PFM)	Public financial management (PFM)	System for the management of public finances; the code unites a set of topics (e.g. budget planning and execution, accounting), thus budget management and budget transparency are more specific sub-categories of PFM, but are often included in general statements on PFM
	Budget management	Management of public spending
	Budget transparency	Openness and accountability related to public spending

4.3 Public financial management

This section summarizes results for the codes on *PFM*, *budget management* and *budget transparency*. The definitions of these codes as used in this evaluation synthesis are stated in Table 21. Most of the codings refer to the broader code PFM as outcome, and both the generic attribution and an input-specific attribution of changes in PFM to non-financial budget support inputs are covered by evidence. This allowed for a differentiated analysis of GBS effects and specific non-financial inputs on PFM. To report the findings in detail, this section is organized by budget support inputs rather than by codes, as in the other sub-sections in this section.

Summary finding: Effects well covered, budget support effective in improving PFM of recipient countries

Of all the codings, the generic attribution of changes in PFM to budget support programmes is best covered by the evidence. The evidence also covers the input-specific effects of non-financial budget support inputs on PFM, although not on budget management and budget transparency. For all three codes, budget support funding effects – also referred to as systemic effects – are insufficiently addressed in the sources.

Budget support programmes are effective in improving the PFM of recipient countries, especially budget formulation and planning, and the comprehensiveness and transparency of the budget. The improvements in PFM are specifically attributed to the non-financial inputs policy dialogue and TA/CD, as priority is generally given to improving PFM

within these inputs. The progress in PFM depends on institutional capacity and political will for PFM reform of the recipient government.

4.3.1 Expected effects

According to the generic intervention logic of budget support, improvements to recipient countries' PFM systems count as one of the key outputs expected to be induced by budget support (see Figure 1). The intervention logic (see section 2.1) draws a direct causal chain from non-financial inputs of budget support programmes to efficient and transparent PFM. In addition, it is assumed that financial inputs also have a direct systemic effect on PFM, as budget support funds are channelled through a country's own budgetary system instead of being channelled through parallel off-budget processes. By using – rather than bypassing – the countries' own budgetary and PFM systems, budget support is thereby expected to create incentives for intrinsic improvements to the countries' PFM systems to make optimal use of these resources. Financial and non-financial inputs are thus expected to reinforce each other in improving PFM and specifically the budget processes in recipient countries.¹³¹

Table 22: Budget support effects on PFM

Inputs		Total	BS Programme	Funds	Policy dialogue	Conditionality	TA / CD
Public Financial Management	# Sources	58	52	9	19	15	29
	Total effect codings	229	122	10	30	17	50
	Positive effect	198	100	8	28	14	48
	No effect	20	13	1	1	3	2
	Negative effect	3	2	0	1	0	0
	Improving effect	0	0	0	0	0	0
	Deteriorating effect	8	7	1	0	0	0
Budget Management	# Sources	34	27	7	7	7	5
	Total effect codings	72	42	8	8	8	6
	Positive effect	61	32	8	8	7	6
	No effect	8	7	0	0	1	0
	Negative effect	3	3	0	0	0	0
	Improving effect	0	0	0	0	0	0
	Deteriorating effect	0	0	0	0	0	0
Budget Transparency	# Sources	29	24	3	3	4	4
	Total effect codings	51	34	4	4	4	5
	Positive effect	40	24	3	4	4	5
	No effect	7	6	1	0	0	0
	Negative effect	2	2	0	0	0	0
	Improving effect	1	1	0	0	0	0
	Deteriorating effect	1	1	0	0	0	0

4.3.2 Budget support programmes in general

Coverage

Among all outcomes, the best coverage is found for the coding on general PFM (see Table 22). Budget support effects on PFM are mentioned by 58 sources, in a total of 229 text sections. Generic effects of budget support are mentioned in 52 sources. Effects on budget management are covered in 34 sources in total, and generic effects of the budget support programmes on budget management are assessed in 27 sources. Effects of budget support on budget transparency are included in 29 sources in total; the generic effects of budget support on budget transparency are covered in 24 sources. This means

that generalizable conclusions can be drawn for the general effects of budget support programmes on PFM, budget management, and budget transparency.

Established evidence

Budget support programmes have a positive effect on PFM in recipient countries. Improvements in PFM are mentioned in 21 sources, reporting improvements in 14 countries^{lxv}.¹³² One of the main reasons for this strong contribution is that budget support enhances the joint focus of donors and the government on the capacity constraints of government performance, in particular in the area of PFM.¹³³ Governments

^{lxv} Burkina Faso, Ghana, Mali, Morocco, Mozambique, Nicaragua, Rwanda, Sierra Leone, Tanzania, Tunisia, Uganda, Vietnam, and Zambia.

Table 23: Specific improvements in PFM

Effect on PFM	Country	Source
Improved budget formulation/ planning	Ethiopia, Ghana, Morocco, Malawi, Mozambique, Rwanda, Tanzania, Uganda, Zambia	Batley et al., 2006; Claussen et al., 2006; Hedger et al., 2010; Lawson et al., 2003b; Ofori/Atta, 2012; Smith, 2009; Steffensen, 2010; Thunnissen/Morillon, 2014; Tidemand, 2009; World Bank, 2008
Improved budgeting/ accounting	Burkina Faso, Ethiopia, Georgia, Malawi, Mozambique, Nicaragua, Rwanda, Sierra Leone, Tanzania, Uganda, Zambia	Claussen et al., 2006; EU, 2014; Gosparini et al., 2006; Hedger et al., 2010; Lawson et al., 2013; Lawson et al., 2016; Leiderer/ Faust 2012; Lister et al., 2006; Purcell et al., 2006; World Bank, 2008; World Bank, 2009a; World Bank, 2013b
Increased comprehensiveness and transparency of the budget	partly Burkina Faso, Burundi, Ghana, Malawi, Mozambique, Nicaragua, Sierra Leone, Tanzania, Tunisia, Uganda, Vietnam	Batley et al., 2006; Bogetic et al., 2015; Claussen et al., 2006; Dijkstra, 2013; DRN 2011; Lanser et al., 2006; Lawson et al., 2007; Lawson et al., 2013; Lawson et al., 2014; Lister et al., 2006; Smith, 2009; World Bank, 2009a; World Bank, 2015b

in countries that receive budget support generally show clear commitment to PFM reform, probably as a result of donor selectivity, given that commitment to PFM reforms usually constitutes an underlying principle on which the provision of budget support is conditioned.

At the generic level, the evidence finds positive effects on budget formulation and planning, budgeting and accounting and comprehensiveness and transparency of the budget (see Table 23). Positive effects on budget formulation and planning are observed in 10 sources covering nine countries, six of them being LICs and three of them being middle-income countries (MICs). Improvements in budgeting and accounting are identified across 12 sources in 11 countries (8 LICs, 3 MICs). The evidence shows generic positive effects on budget transparency (13 sources; see also section 4.5 on domestic accountability).

Overall, the sources draw a distinct positive picture on budget support's influence on PFM (see Table 22), and only a few sources report negative effects. These sources discuss negative effects of budget support suspensions on fiscal discipline and a decline in performance in PFM, in particular regarding budget credibility and transparency¹³⁴ (see paragraphs below).

In some country case examples,¹³⁵ the introduction or further roll-out of a financial management information system,¹³⁵ the introduction of laws and regulations regarding public finance and budget management, or the enhancement of existing legal

frameworks led to improvements in PFM.¹³⁶ Regarding institutions of accountability, the Office of the Auditor General and the role of Parliament in the budget process are strengthened in some cases (see also section 4.5 on domestic accountability).¹³⁷ In other cases, budget support resulted in improvements of procurement processes,¹³⁸ more reporting (also internal and external auditing) and control over government expenditure.¹³⁹

In spite of the distinctly positive effects of budget support programmes on PFM, the budget support effects on PFM reforms did not live up to expectations. The underachievement of PFM reform is among other reasons explained by weak capacities and weak institutional structures undermining stronger progress, which is mentioned in seven country cases.¹⁴⁰ Furthermore, in many countries reforms have yet to be implemented and budget information provided at local government levels as the reforms focused so far on improvements at central level.¹⁴¹

Compared over the life-span of budget support programmes, the initial momentum to implement PFM reform was lost as a slow-down or delays in PFM reforms set in towards the later years of budget support programmes, for example in Ethiopia, Ghana, Vietnam, and Zambia.¹⁴² Explanations provided in the sources are the diminution of reform commitment coinciding with the reduction of budget support contributions, and the long maturation process needed for the implementation of PFM reforms.¹⁴³

¹³⁴ Ghana, Malawi, Mozambique, Nicaragua, Sierra Leone, Uganda, and Zambia.

¹³⁵ Ghana, Sierra Leone, Uganda, and Tanzania.

¹³⁶ Ghana, Morocco, Mozambique, Nicaragua, Sierra Leone, Tanzania, Uganda, and Vietnam.

¹³⁷ Burkina Faso, Malawi, Mali, Nicaragua, Sierra Leone, Tanzania, and Zambia.

¹³⁸ e.g. Burkina Faso, and Uganda.

4.3.3 Financial and non-financial inputs

Coverage

An effect of TA/CD on PFM is described in 29 sources, an effect of policy dialogue in 19 sources, and 15 sources state an effect of conditionality on PFM. There is insufficient evidence on effects of funding on PFM, which is only covered in nine sources. The insufficient coverage of funding effects on PFM means that generalizable conclusions on effects, as laid out in the intervention logic,¹⁴⁴ such as the systemic effect of budget support, cannot be drawn from the evidence.

As shown in Table 22, for both budget management and budget transparency the coverage of the individual inputs is insufficient, and therefore no generalizable conclusions with respect to effects of individual inputs on budget management and budget transparency can be drawn. As the sources do not clearly differentiate between the codes PFM, budget management and budget transparency, some findings on the effect of non-financial inputs on budget management and transparency are subsumed under the topic PFM, which is more broadly used in the sources.

Established evidence

Overall, the non-financial inputs of budget support have a positive effect on improvements in PFM.¹⁴⁵ This is reported in 19 sources.¹⁴⁶ Two of these reports attribute achievements in PFM mainly to non-financial inputs of budget support, implying that it was primarily non-financial rather than financial budget support inputs that had a positive effect on PFM, due to increased awareness for reform needs in the field of PFM and the supply of inputs for the PFM reform process.¹⁴⁷ This finding supports the assumption of the budget support intervention logic, that the combination of financial and non-financial inputs reinforces positive effects on PFM.

TA/CD as a non-financial input of budget support programmes (see Box 4) has proven positive effects on PFM, as it addresses capacity needs in PFM and governance institutions (attested in 10 sources).¹⁴⁸ TA/CD improved, for example, both planning and budgeting in six countries,¹⁴⁹ by helping build capacities on an institutional level at the Office of the Auditor

General (OAG) and offering technical assistance in budget management and other PFM functions.¹⁵⁰ One of the few criticisms with regard to TA/CD accompanying budget support is the lack of coordination in the technical assistance provided by donors, both within budget support programmes and in the interaction of budget support programmes, with TA/CD offered in other programmes, for example in Uganda and Zambia.¹⁵¹

Box 4 Evidence on accompanying TA/CD

Although the effects of TA/CD on PFM are positive, evidence is predominantly found in more recent evaluation reports. Most evaluations applied a narrow definition, and counted TA/CD only as part of the budget support package if it is part of the same financing agreement; it therefore analysed effects of TA/CD only to a limited extent. In practice, however, donors accompany their budget support programmes with parallel TA/CD and consider these accompanying measures as part of the budget support package. Based on a broader definition of accompanying measures (all forms of TA/CD that are provided simultaneously to budget support and pursue related objectives), Krisch et al. (2015) emphasize the relevance of TA/CD, and more recent evaluations recognize the contributions of TA/CD, as an important element of budget support, in directly strengthening national systems, particularly those of PFM.

Regarding policy dialogue, the sources included in the evaluation synthesis attest a positive effect on PFM, in particular on budget management¹⁵² and budget transparency.¹⁵³ The dialogue works in favour of PFM, as PFM and related PAF indicators are prioritized, such as alignment of funds towards national strategies and policy frameworks.¹⁵³ While the policy dialogue initiated reforms in most cases,¹⁵⁴ it also helped to keep the recipient government on track in the follow up of reforms.¹⁵⁴

Although policy dialogue overall has a positive effect on PFM,

¹⁴⁴ Mozambique, Sierra Leone, Tanzania, Tunisia, Uganda, and Zambia.

¹⁴⁵ e.g. in Mozambique, Nicaragua, and Rwanda.

¹⁴⁶ e.g. in Nicaragua and Zambia.

¹⁴⁷ E.g. Mali, Morocco, Rwanda, South Africa, Tanzania, Tunisia, Uganda, and Vietnam.

¹⁴⁸ E.g. in Burkina Faso, Ghana, Mozambique, Nicaragua, and Sierra Leone.

Table 24: Corruption codes

Topic	Codes	Definition
Corruption	Anti-corruption agency	Agency investigating against corruption cases
	Corruption legislation	Regulations to fight corruption
	Corruption level	Degree of corruption
	Corruption policies	Practices and rules regarding corruption

some sources argue that budget support dialogue does not use its potential to the fullest. Only modest effects of budget support dialogue on PFM are visible in areas where the political support for reforms is lacking, or where changes in the priorities of socio-economic policy are necessary.¹⁵⁵ This is also attributed to the conflictive nature of dialogue in some countries and the absence of joint dialogue strategies.¹⁵⁶

The sources state that conditionality in the form of PAF and diagnostic tools helps to improve PFM, as the focus of donors and government shifts towards improvements in PFM indicators (a large share of PAF indicators dealt with improvements in PFM¹⁵⁷) and constant review processes on the performance in PFM provide incentives to persistently implement PFM reforms.¹⁵⁸ With regard to budget transparency, the PAF positively influences the traceability of government processes and expenditures.¹⁵⁹ However, whether PFM criteria in the PAF are conducive to functional PFM changes seems to depend highly on the willingness to reform.¹⁶⁰ Moreover, in some countries it is assumed that PFM capacity constraints were aggravated by insufficient harmonization of procedures and extensive control requirements.¹⁶¹

4.4 Corruption

The topic corruption is divided into the codes *corruption level*, *corruption policies*, *corruption legislation* and *anti-corruption agencies*. For the specific definitions of the codes as used in this evaluation synthesis see Table 24.

Summary finding: Surprisingly low coverage on highly debated effects on corruption

Evidence on budget support's effects on corruption is scarce in the sources. Considering that two out of four codes on corruption do not meet the threshold of coverage in 10 reports, and only a low number of codings is available, most effects on corruption are insufficiently backed by evidence. This lack of evidence on corruption is both surprising and problematic as the political debate revolves to a large extent around budget support's effects on corruption.

However, the analysed data show no negative effect of budget support on corruption and thus do not confirm this assumed risk. One effect that is sufficiently covered by evidence relates to the monitoring of corruption, which increases due to budget support. However, the prosecution of people accused of corruption is still low and the number of convictions even lower.

4.4.1 Expected effects

Taking a certain level of corruption as a given, the political debate controversially discusses the fiduciary risks connected to the disbursements of budget support. The budget support intervention logic, on the other hand, assumes that increases in transparency and accountability due to budget support contribute to a reduction in corruption.

Table 25: Budget support effects on corruption

Inputs		Total	BS Programme	Funds	Policy dialogue	Conditionality	TA / CD
Anti-corruption Agency	# Sources	9	9	2	1	1	2
	Total effect codings	19	13	2	1	1	2
	Positive effect	17	12	1	1	1	2
	No effect	2	1	1	0	0	0
	Negative effect	0	0	0	0	0	0
	Improving effect	0	0	0	0	0	0
	Deteriorating effect	0	0	0	0	0	0
Corruption Legislation	# Sources	5	5	0	0	0	0
	Total effect codings	5	5	0	0	0	0
	Positive effect	5	5	0	0	0	0
	No effect	0	0	0	0	0	0
	Negative effect	0	0	0	0	0	0
	Improving effect	0	0	0	0	0	0
	Deteriorating effect	0	0	0	0	0	0
Corruption Level	# Sources	19	15	3	1	2	1
	Total effect codings	25	17	3	1	3	1
	Positive effect	10	8	1	1	0	0
	No effect	13	7	2	0	3	1
	Negative effect	1	1	0	0	0	0
	Improving effect	1	1	0	0	0	0
	Deteriorating effect	0	0	0	0	0	0
Corruption Policies	# Sources	13	10	0	2	3	1
	Total effect codings	19	11	0	3	4	1
	Positive effect	11	6	0	3	2	0
	No effect	8	5	0	0	2	1
	Negative effect	0	0	0	0	0	0
	Improving effect	0	0	0	0	0	0
	Deteriorating effect	0	0	0	0	0	0

4.4.2 Effects on corruption

Coverage

Out of 95 sources, effects on corruption levels are covered in a total of 19 sources, effects on corruption policies in 13 sources, effects on corruption legislation in five sources, and effects on anti-corruption agencies in nine sources (see Table 25). The number of codings is either the same or only slightly higher than the number of reports covering the effect, which implies that none of the sources extensively cover effects on corruption. None of the effects of specific inputs on any of the corruption codes reaches the threshold for empirically established effects, thus no conclusions can be drawn on the impact of financial and non-financial inputs on corruption. Whether an increase in transparency and accountability due to budget support leads to a reduction in corruption cannot be answered based on these findings. Future research is needed to close the knowledge gap regarding budget support effects on corruption.

Established evidence

Due to the low number of codings on all corruption codes (see Table 25), very few effects on corruption meet the threshold to be included as established evidence. Overall, the number of codings presented in Table 25 suggests that there is no negative effect of budget support on the corruption level. The evidence is mixed regarding the question of whether there is no effect of budget support on corruption or whether there is a positive one. The one positive effect that is addressed in more than 10 reports is the improvement in monitoring of corruption due to budget support (12 sources).¹⁶² This stems from positive budget support effects on domestic accountability and PFM. In particular, TA/CD as budget support input strengthened institutions and stakeholders to detect corruption, mainly anti-corruption agencies. However, the prosecution of people accused of corruption is still low and the number of convictions even lower (4 sources).¹⁶³

For corruption policies, budget support clearly has no negative effect. Once again, evidence is mixed regarding no or positive budget support effects, although corruption is formally on the agenda of the (high level) policy dialogue and directs the partner governments' attention towards their anti-corruption policies.¹⁶⁴

4.5 Domestic accountability and democratic governance

The topic domestic accountability and democratic governance is sub-divided into the supply side of domestic accountability, the demand side of domestic accountability and democratic governance. Institutions on the supply side of domestic accountability include government institutions like the ministry of finance and planning, national statistic offices, and supreme audit institutions (SAI),^{lxxxvi} expected to supply (budget) information to the general public. The demand side of domestic accountability consists of actors, like the parliament, civil society and media, who use this information to hold the government to account. These sub-topics define the structure of the following paragraphs. Table 26 indicates the specific definitions of the codes used in this evaluation synthesis.

Summary finding: Budget support improves the supply side of domestic accountability and strengthens SAI on a generic level, no input-specific attribution

The generic attribution of changes in domestic accountability to budget support programmes is sufficiently covered by evidence, with the highest coverage for domestic accountability and the lowest coverage for SAI. However, none of the codes has sufficient coverage on financial and non-financial inputs and their effects to draw generalizable conclusions.

Budget support contributes to improvements on the supply side of domestic accountability. Budget support strengthens the budget process as it is channelled through the national budget system. It substantially augments the role of SAI, where improvements are found, particularly regarding the quality and quantity of audit reports. With regard to the demand side of accountability, long-term effects of budget support on the role of parliament and civil society could not be identified.

^{lxxxvi} Note that SAI can be part of the supply as well as the demand side and fulfil both functions depending on the institutional and legal status.

Table 26: Domestic accountability and democratic governance codes

Topic	Sub-topic	Codes	Definition
Domestic Accountability and Democratic Governance	Supply side	Domestic accountability	Governmental liability related to public spending
		Supreme audit institutions	National agencies responsible for auditing government revenue and spending and their strengthened function
	Demand side	Role of Parliament	Strengthened function of parliament
		Civil society organizations	Strengthened function of non-governmental organizations
	Democratic governance	Governance	Relates to areas of governance that are not covered under domestic accountability, such as rule of law, intra-government accountability, decentralization

4.5.1 Expected effects

As per the CEF, governance and accountability objectives are at the core of the intervention logic of budget support. Besides their aim to reduce poverty and promote socio-economic development, donors set good governance as an ultimate objective of budget support. The CEF describes improvements in human rights and democracy at the impact level. Furthermore, good governance is not only an end in itself but serves as a cross-cutting issue to increase the effectiveness of budget support to achieve its objectives.¹⁶⁵ On the intermediary level of induced outputs, budget support is expected to positively affect the supply and the demand side of domestic accountability. Expected effects include strengthened public sector institutions of PFM, improved budget transparency and strengthened linkages between government and oversight bodies¹⁶⁶.

4.5.2 Supply side of domestic accountability

Coverage

Budget support effects on domestic accountability are covered in 44 sources and in a total of 123 codings, which is the highest coverage among all codes on domestic accountability (see Table 27). Statements regarding the generic attribution of changes in domestic accountability to budget support programmes are found in 42 sources. The input-specific effects of funding (8 sources), policy dialogue (9 sources), conditionality (7 sources) and TA/CD (8 sources) are all insufficiently covered in the body of evidence with respect to the benchmark of 10 sources. Effects of budget support on SAI are covered in 17 sources, with 15 sources covering the generic effect of budget support programmes, and fewer than 10 sources describing

effects from individual inputs of budget support. This means that no conclusions can be drawn on the effects of specific budget support inputs on domestic accountability of SAI.

Established evidence

The overall finding that emerges from synthesizing the reviewed evidence is that budget support generally contributes to improvements on the supply side of domestic accountability. Positive effects have been identified in nine countries.¹⁶⁷ The effects of budget support on the supply side of domestic accountability include significant effects on PFM, particularly the increased comprehensiveness and transparency of the budget (see section 4.3). Interestingly, observed changes in domestic accountability have usually not been attributed to one or a combination of particular inputs of budget support. Instead, the great majority of statements describe aggregate effects of the budget support programme more generally.

Positive effects of all budget support inputs on the supply side of domestic accountability are reported in 38 sources.¹⁶⁸ As budget support is channelled through the national budget systems, it further increases the relevance of these systems. National processes in the budget cycle are backed by budget support because external flows are managed by the Ministry of Finance and are subject to decision making by Cabinet and approval by Parliament.¹⁶⁹

There is sufficient coverage on the link between budget support and SAI. Of the 17 sources covering effects on SAI, 14 report positive effects.¹⁷⁰ The most significant improvements

¹⁶⁵¹⁶⁶¹⁶⁷¹⁶⁸¹⁶⁹¹⁷⁰ Burkina Faso, Ghana, Mozambique, Rwanda, Tanzania, Tunisia, Uganda, Vietnam, and Zambia.

Table 27: Budget support effects on the supply side of domestic accountability

Inputs	Total	BS Programme	Funds	Policy dialogue	Conditionality	TA / CD
Domestic Accountability	# Sources	44	42	8	9	8
	Total effect codings	123	88	9	9	8
	Positive effect	100	71	5	9	7
	No effect	19	14	4	0	1
	Negative effect	3	2	0	0	0
	Improving effect	1	1	0	0	0
	Deteriorating effect	0	0	0	0	0
Supreme Audit Institutions	# Sources	17	15	6	2	4
	Total effect codings	34	18	6	2	4
	Positive effect	29	14	6	2	3
	No effect	4	3	0	0	1
	Negative effect	1	1	0	0	0
	Improving effect	0	0	0	0	0
	Deteriorating effect	0	0	0	0	0

have been identified regarding the quality and quantity of audit sources as well as their timely publication. However, the evidence suggests that achievements have been limited to the technical level, and budget support was not able to promote solutions related to political issues such as the follow up and implementation of audit recommendations, the enforcement of prosecution, and the recovery of funds.¹⁷¹

Ten sources point to the potentially conflicting relation between the ultimate objective to strengthen domestic accountability and the external accountability demands from donors vis-à-vis the recipient government.¹⁷² In fact, two extreme examples that point to adverse accountability effects are Burundi, where the budget support donor group took over accountability functions from local institutions, and Uganda, where donors dominated the dialogue at the expense of domestic stakeholders.¹⁷³ At the same time, there are also some examples where accountability systems improved in a way that is of value to both domestic and international stakeholders, and where donors' accountability demands helped fill a gap in the recipient country's monitoring and review processes.¹⁷⁴ However, key for achieving such win-win

situations is the use of policy dialogue as a means to raise domestic awareness and knowledge of key policy debates and to publicly share all information in order to allow domestic accountability institutions to follow up.¹⁷⁵

4.5.3 Demand side of domestic accountability

Coverage

Evidence regarding the demand side of domestic accountability and democratic governance is limited.¹⁷⁶ The effects of budget support on the role of parliament are covered in 21 sources, of which 20 cover the generic effects of budget support programmes. The effects of budget support on civil society are covered by 25 sources, with 24 covering the generic effects of budget support programmes. As displayed in Table 28, individual inputs of budget support are insufficiently covered by evidence. This means that no generalizable conclusions can be drawn on the specific effect of each input on the demand side of domestic accountability. For example, it cannot be assessed whether the assumption that non-financial inputs strengthen civil society engagement and the oversight function of parliament applies in reality.

Table 28: Budget support effects on the demand side of domestic accountability

Inputs		Total	BS Programme	Funds	Policy dialogue	Conditionality	TA / CD
Role of Parliament	# Sources	21	20	4	4	1	2
	Total effect codings	56	39	6	5	4	2
	Positive effect	38	27	5	3	2	1
	No effect	18	12	1	2	2	1
	Negative effect	0	0	0	0	0	0
	Improving effect	0	0	0	0	0	0
	Deteriorating effect	0	0	0	0	0	0
Civil Society Organizations	# Sources	25	24	4	8	1	2
	Total effect codings	68	51	6	8	1	2
	Positive effect	51	36	6	7	1	1
	No effect	14	12	0	1	0	1
	Negative effect	2	2	0	0	0	0
	Improving effect	0	0	0	0	0	0
	Deteriorating effect	1	1	0	0	0	0

Established evidence

The effects of budget support in strengthening the oversight function of parliaments are covered by 21 sources, of which 14 indicate a positive effect. However, only two of the codings indicate a strong positive effect, and the majority of the codings (18) indicate a weak positive effect. Nine sources find no effect.¹⁷⁷ Moreover, some evaluations and studies make statements on (changes of) the role of parliaments in the budget process without sufficiently attributing these to the presence of the budget support programme. The main barriers to effective budget oversight by parliament identified in the literature are limited capacities, lack of information, and political constraints.¹⁷⁸

Results are even more inconsistent for the effect on the participation of civil society organizations (CSO) in the budget process. The quantitative analysis results in 18 sources that report positive effects.¹⁷⁹ However, the qualitative assessment of the text segments reveals that statements mostly describe marginal improvements without attributing the observed changes to the budget support programme. The majority of text segments have thus been taken as weak positive effects.

Another 10 sources state absent or negative effects.¹⁸⁰ In some country cases^{lxxxviii} the evidence indicates that CSOs have not been sufficiently integrated.¹⁸¹ However, in Malawi and Mozambique a dialogue between the Ministry of Finance, the sector ministries, and the National Statistical Office (Malawi)/ National Institute of Statistics (Mozambique) was established. CSOs participated through the Poverty Observatory.¹⁸² This “inclusive” budget support dialogue helped to create political space for CSOs and the media to participate and address issues of political and financial accountability.¹⁸³ Despite improvements in CSO participation in budget processes described in these and other country case studies,¹⁸⁴ more critical studies voice concerns on the ability of CSOs to effectively hold the government to account.¹⁸⁵

Overall, budget support seems to perform below its potential to effectively support the demand side of domestic accountability. The failure to directly support non-state actors in the budget process has been linked to a lack of focus of non-financial inputs (e.g. policy dialogue, conditionality, and TA/CD) of budget support (for example, in Mozambique and Zambia).¹⁸⁶ Moreover, the induced improvements on the side

^{lxxxviii} Rwanda, Uganda, Vietnam, and Zambia.

Table 29: Budget support effects on governance

Inputs		Total	BS Programme	Funds	Policy dialogue	Conditionality	TA / CD
Governance	# Sources	31	27	1	8	6	4
	Total effect codings	65	41	1	8	8	7
	Positive effect	53	33	1	7	6	6
	No effect	11	7	0	1	2	1
	Negative effect	0	0	0	0	0	0
	Improving effect	0	0	0	0	0	0
	Deteriorating effect	1	1	0	0	0	0

of SAI and, to a lesser extent, also on the side of parliament and CSOs have been largely of a technical nature, while progress towards a broader governance objective has not been realized.¹⁸⁷ Donors did not succeed in changing the recipient government's priorities towards the demand side of accountability or to improve democratic governance more generally.¹⁸⁸

4.5.4 Democratic governance

Coverage

The effects of budget support on governance are covered in 31 sources, of which 27 deal with generic effects of budget support. In line with the other codes on domestic accountability, the effects of individual inputs on governance are not sufficiently covered by evidence (see Table 29). This means that no generalizable conclusions can be drawn on the effects of specific inputs on governance, specifically regarding the expected effect of policy dialogue on improving governance through better policy choices.¹⁸⁹

Established evidence

The majority of sources describe positive effects of budget support on governance,¹⁹⁰ and few sources report absent effects.¹⁹¹ However, the effects on governance are not clear cut, as suggested by the quantitative results displayed in Table 29 above. The qualitative assessment of the 65 codings reveals a more nuanced picture. Most sources only draw hesitant conclusions on the effects of budget support on good governance. Statements generally describe "limited" or "mild" effects, without providing details on the specific contribution

of the budget support programme. The codings include broad statements and cover different aspects of governance. The improvement in governance particularly refers to stronger intra-government accountability structures.¹⁹² In contrast to conventional project aid, which is typically implemented via specific sector ministries, budget support disbursements are allocated by the Ministry of Finance. In Uganda and Mozambique, the additional on-budget resources provided through budget support contributed to strengthening the accountability relation between the Ministry of Finance, sector ministries, and the national statistical office.¹⁹³ Moreover, instead of turning to the donor community, sector ministries, in order to secure funding, stick to the national budget process.¹⁹⁴ Uncertainty remains with regard to the effect of budget support on decentralization. While budget support facilitated decentralization in some countries (Mali, Mozambique, South Africa, Uganda (in earlier years), Tanzania, and Vietnam),¹⁹⁵ the instrument did not achieve progress in decentralization in Nicaragua, Rwanda, Tunisia, and Uganda (in later years).¹⁹⁶

Table 30: Economic performance and income poverty codes

Topic	Code	Definition
Economic Performance and Income Poverty	Macroeconomic performance	Macroeconomic management that improves economic growth and stability
	Income poverty	Disposable income of below the poverty line

4.6 Economic performance and income poverty

This section analyses evidence on macroeconomic performance and income poverty. The definitions of the codes as used in this evaluation synthesis are stated in Table 30.

Summary finding: Pre-existing macroeconomic stability reinforced, but causal chain from economic growth to less income poverty not supported (problems of attribution)

The generic attribution of changes in macroeconomic performance and income poverty to budget support programmes is sufficiently covered by evidence. Concerning specific budget support inputs, only the effect of funding on income poverty is sufficiently covered, for all other inputs, no generalizable conclusions on effects can be drawn.

While positive developments in macroeconomic performance are not attributed to the presence of budget support programmes alone, the evidence finds that budget support programmes reinforce pre-existing macroeconomic stability. A certain repayment of domestic debt as one trigger of economic growth is also plausibly attributed to budget support programmes. The evidence on budget support effects on income poverty is mixed. Most sources find no or only slightly positive effects, but reliable evidence on the contribution of budget support to these changes is relatively scarce.

4.6.1 Expected effects

Based on the intervention logic, two major effects are expected from budget support programmes on economic performance and income poverty: a systemic effect and a policy and institutional effect. The systemic effect is expected to allow the government an increase in public spending or

saving, or a reduction in public borrowing (see Section 4.2). The budget support funding might in turn also have macroeconomic effects, particularly on economic growth, and on interest and exchange rates. The assumption is that budget support creates added demand for domestic products and allows a reduction in government borrowing. Second, policy and institutional effects result from non-financial budget support inputs via improved macroeconomic and sector policy management.¹⁹⁷ The progress in economic development due to budget support is expected to result in a reduction of income poverty.¹⁹⁸ However, this is the product of the entirety of government policies, decisions on spending, and budget allocation, and is influenced by budget support only to a limited degree.¹⁹⁹

4.6.2 Macroeconomic performance

Coverage

The effects of budget support on macroeconomic performance are covered in 31 sources, of which 28 cover effects of budget support programmes at a generic level. Effects of the individual budget support inputs funding, policy dialogue, conditionality and TA/CD on macroeconomic performance are covered in an insufficient number of sources (see Table 31). The evidence base is insufficient for conclusions to be drawn on the effectiveness of specific budget support inputs as formulated in the intervention logic, e.g. the expected positive effect of budget support funding on, for example, the balance of payments, savings and macroeconomic stabilization²⁰⁰.

Established evidence

Overall, the evidence describes positive effects on macroeconomic performance. These effects are reported for 17 countries,^{lxxxix 201} of which 13 are LICs.²⁰² However, most sources agree that the evidence base is insufficient to clearly attribute improvements in macroeconomic performance to the presence of budget support programmes.²⁰³ Many sources do

^{lxxxix} Afghanistan, Burkina Faso, Burundi, CAR, DRC, Ethiopia, Ghana, Malawi, Mali, Mozambique, Nepal, Nicaragua, Sierra Leone, Tanzania, Tunisia, Uganda, and Zambia.

Table 31: Budget support effects on macroeconomic performance

Inputs		Total	BS Programme	Funds	Policy dialogue	Conditionality	TA / CD
Macroeconomic Performance	# Sources	31	28	8	3	4	1
	Total effect codings	65	43	14	3	4	1
	Positive effect	56	36	14	3	2	1
	No effect	1	1	0	0	0	0
	Negative effect	6	4	0	0	2	0
	Improving effect	0	0	0	0	0	0
	Deteriorating effect	2	2	0	0	0	0

not attribute improvements in economic performance to budget support alone, but also, sometimes more prominently, to programmes of IMF and to the political commitment of the partner governments to achieve macroeconomic stability (particularly in the form of monetarist policies).²⁰⁴ What the evidence does attribute to the presence of budget support is that budget support programmes reinforce already existing macroeconomic stability, as the programmes involve a higher number of international partners in the macroeconomic debate, and foster fiscal discipline by the provision of on-budget funds.²⁰⁵

Growth of gross domestic product is reported in nine countries^{xc}, four of them LICs and five MICs,²⁰⁶ but the attribution to budget support programmes is inconsistent across the sources. Some sources suggest that budget support is supportive of economic growth, as budget support funds enable increases in government spending without a simultaneous increase in domestic borrowing.²⁰⁷ Other sources find that economic growth is predominantly caused by debt relief, high domestic commodity prices and IMF programmes.²⁰⁸ The repayment of domestic debt as one trigger of economic growth is plausibly attributed to budget support programmes. This effect is found in six sources²⁰⁹ and in five country cases.^{xc1} The underlying rationale is that budget support enabled an increase in internal repayment of the budget deficit. This argument is supported by the immediate increase in domestic borrowing as a result of budget support suspensions.²¹⁰

While the link between increased economic performance and budget support programmes is questionable, a few sources suggest the negative impact of budget support suspensions on economic performance. Coinciding with low domestic-revenue mobilization, mixed and even adverse effects on macroeconomic performance, such as fiscal indiscipline, macroeconomic instability and increases in budget deficit, are reported.²¹¹

Budget support has a weak positive effect on private sector development, especially on the business environment. The reported effects range from an assessment as “unsatisfactory” in Burundi, Ethiopia, and Mozambique, through weak positive effects in Burkina Faso, Ghana, and Santa Lucia, to strong positive effects in South Africa and Tunisia. The relevance of budget support effects on private sector development in policy and aid strategies is mentioned in early evaluation sources, but generally seems to be gaining in importance.²¹²

^{xc} Ghana, Mozambique, Mali, Nicaragua, Sierra Leone, Tanzania, Tunisia, Vietnam, and Zambia

^{xc1} Ghana, Nicaragua, Sierra Leone, Tanzania, and Zambia.

Table 32: Budget support effects on income poverty

Inputs		Total	BS Programme	Funds	Policy dialogue	Conditionality	TA / CD
Income Poverty	# Sources	40	36	17	7	4	1
	Total effect codings	127	87	24	8	7	1
	Positive effect	77	51	15	6	4	1
	No effect	38	29	6	2	1	0
	Negative effect	5	4	0	0	1	0
	Improving effect	3	2	1	0	0	0
	Deteriorating effect	4	1	2	0	1	0

4.6.3 Income poverty

Coverage

The coverage on budget support effects on income poverty in the sources is strong (in 40 out of 95 sources)²¹³ (see Table 32). The generic effect of budget support is analysed in 36 sources, and 17 sources analyse the input-specific effect of budget support funding on income poverty. The effect of non-financial inputs is insufficiently covered by evidence (seven sources for policy dialogue, four sources for conditionality and one source for TA/CD). Thus, no generalizable conclusions on the effects of individual non-financial inputs on income poverty can be drawn.

Established evidence

As stated in Section 4.6.1, the increase in economic performance due to budget support is expected to result in a reduction of income poverty. While increases in economic performance and growth are confirmed by the evidence, the generic attribution of these improvements to budget support programmes is not in all cases plausibly confirmed or disconfirmed by the evidence.

The attribution of changes in income poverty to budget support programmes is particularly difficult as budget support promotes general reforms, which only indirectly enhance poverty reduction. Moreover, in some country cases^{xcii} reductions in income poverty were not visible at the time of evaluation due to time lags between cause and effect. This fundamental challenge is mostly reported in more recent sources.²¹⁴ The causal link between budget support and income

poverty is hence strongly diluted, which is reflected in the effects presented by the evidence.

The body of evidence, overall, finds inconsistent effects on income poverty:²¹⁵ the majority of sources report either no or slightly positive effects of budget support on income poverty. No effect of budget support programmes on income poverty is reported in seven countries.^{xciii} ²¹⁶ In contrast, positive effects of budget support on income poverty are found in eight sources,²¹⁷ covering seven countries.^{xciv}

While the attribution to budget support remains unclear, the evidence suggests that the effects on income poverty did not develop as expected, and offers a variety of reasons why this is the case. Most of these reasons reported are specific to the respective country context, but some are found in several cases. One reason mentioned is a poor conceptualization of pro-poor policies (Tanzania and Zambia)²¹⁸, which neglected inequalities in the poor's access to assets (Zambia, Mozambique, Uganda, and Nicaragua).²¹⁹ The evidence also refers to policy mistakes, particularly in the agriculture sector, such as subsidies and programmes that did not target poor farmers. Overall, macroeconomic growth did not reach the relevant population groups, particularly in rural areas.²²⁰

^{xcii} Burkina Faso, Ghana, Malawi, Mozambique, Nicaragua, Rwanda, Uganda, and Vietnam.

^{xciiii} Burundi, Ghana, Malawi, Mozambique, Nicaragua (under the Bolaños government), Rwanda, and Sierra Leone (in the Western region).

^{xciv} Mali, Nicaragua (under Ortega government), Sierra Leone (Eastern and northern regions), South Africa, Uganda, and Zambia.

Table 33: Service delivery and non-income poverty codes

Topic	Code	Definition
Service Delivery and Non-Income Poverty	Public service delivery	Provision of public services
	Administration for service delivery	Managing the provision of public services
	Non-income poverty	Lack of access to goods and amenities that are deemed to be needed for an acceptable standard of living

4.7 Service delivery and non-income poverty

The subsequent sections on service delivery and non-income poverty summarize the evidence from codings on *overall public service delivery*, *public administration for service delivery* and *non-income poverty*. The specific definitions of these codes as used in this evaluation synthesis are stated in Table 33.

Summary finding: Access to public services increased, but of low quality that impairs the effect on non-income poverty – problems of attribution

The generic attribution of changes in service delivery and non-income poverty to budget support is sufficiently covered by evidence, with the highest coverage for public service delivery and the lowest coverage for administration for service delivery. Regarding the effects of specific budget support inputs, the effects of funding, policy dialogue and TA/CD on public service delivery are also sufficiently covered, though conditionality is not. For administration of public service delivery and non-income poverty, none of the specific inputs has sufficient coverage in the sources.

Overall, budget support is effective at increasing access to public services. Funds from SBS and GBS are decisive in increasing service delivery, particularly in the education and health sectors. Yet, budget support does not prove to be effective at improving the quality of provided services or the administration responsible for delivering the services. The evidence describes positive effects on non-income poverty in the recipient countries, but the effects are not attributed to budget support alone.

4.7.1 Expected effects

Outcomes expected from budget support are improvements in public service delivery and the reduction of non-income poverty. By providing financial means to the recipient government's budget, increased expenditure is expected to be used to finance the expansion of service delivery in the social sectors and is supposed to lead to reduced levels of non-income poverty.^{xv} Through its non-financial elements, budget support is also designed to influence sector policies and increase the quality of social services by strengthening the administrative capacities for service delivery.

4.7.2 Public service delivery

Coverage

Effects of budget support on public service delivery are examined in 47 sources and are mentioned in a total of 186 text sections, which means that public service delivery has the second strongest coverage in the entire body of evidence (see Table 34). Generic effects of budget support programmes are covered in 40 sources, and 28 cover effects relating to the financing function of budget support. Effects of policy dialogue and TA/CD are both covered in 10 sources. Effects of budget support on conditionality are, however, insufficiently covered, with seven sources, and thus no generalizable conclusions on the effects of this specific input can be drawn.

The coverage for public administration for service delivery is lower, but still sufficient, with 25 sources on the effects of budget support. Generic effects of budget support are covered in 18 sources, but there is only weak coverage for specific inputs of budget support. Funding and TA/CD are covered in six sources and policy dialogue and conditionality are covered in five sources. Specific inputs of budget support are thus insufficiently covered to derive generalizable conclusions.

^{xv} Non-income poverty is defined by the level of social exclusion and the lack of education, health, and social welfare (Lawson et al., 2013; Lawson et al., 2014; Lawson et al., 2016). The dimensions of non-income poverty most commonly covered in evaluations are education and health (through indicators from the Millennium Development Goals (e.g. mentioned in Lister, 2006), or the HDI (Lawson et al., 2013)).

Table 34: Budget support effects on service delivery

Inputs		Total	BS Programme	Funds	Policy dialogue	Conditionality	TA / CD
Public Service Delivery	# Sources	47	40	28	10	7	10
	Total effect codings	186	106	43	15	9	13
	Positive effect	150	94	35	11	3	7
	No effect	33	12	6	4	5	6
	Negative effect	2	0	1	0	1	0
	Improving effect	0	0	0	0	0	0
	Deteriorating effect	1	0	1	0	0	0
Administration for Service Delivery	# Sources	25	18	6	5	5	6
	Total effect codings	50	27	6	6	5	6
	Positive effect	44	24	6	5	4	5
	No effect	4	3	0	0	0	1
	Negative effect	2	0	0	1	1	0
	Improving effect	0	0	0	0	0	0
	Deteriorating effect	0	0	0	0	0	0

Established evidence

There is a large body of evidence on the effect of budget support on public service delivery (47 sources). A majority of these sources provide evidence for a positive effect of budget support on public service delivery, which has been identified in 44 sources. Some sources report that it was not possible to assess the instrument's effects on service delivery because effects were not yet perceptible in the country cases at the time of evaluation.²²¹ The evidence suggests that funds provided through budget support (together with domestic resources) are predominantly used for investment or to cover running costs in the social sectors (e.g. salaries for health workers or teachers) and thus increase people's access to basic services. However, due to fungibility of funds, the sources cannot state with absolute certainty for which purposes budget support disbursements were used.

Substantial improvements in the outreach of public services are noted, particularly in the sectors of education (10 countries^{xcvi} in 19 sources)²²² and health (11 countries^{xcvii} in 18 sources).²²³

The effect of budget support on administration of service delivery is covered in 25 sources,²²⁴ but the direction of the effect is inconsistent. While improvements are described in eight countries,^{xcviii} in four countries^{xcix} no effects on capacity and administration for service delivery are found.²²⁵

Despite substantial improvements in people's access to public services due to budget support, the evidence does not provide clear results on the effects of budget support on the quality of provided services. In ten countries where budget support contributes to expanding the scope of public services, 14 sources find stagnating levels of low quality in basic services.²²⁶

More recent evaluations, while identifying further increases in coverage of service delivery and positive trends in overall literacy rates,²²⁷ identify trade-offs and failures of budget support to effectively address issues of quality in service delivery. A trade-off is the priority given to schooling access over quality and efficiency in primary education. A failure is the neglect of service management and delivery issues at

^{xcvi} Burkina Faso, Ghana, Mali, Morocco, Mozambique, Nicaragua, Sierra Leone, Tanzania, Uganda, and Zambia.

^{xcvii} Burkina Faso, Ghana, Mali, Morocco, Mozambique, Nicaragua, Sierra Leone, South Africa, Tanzania, Uganda, and Zambia.

^{xcviii} Ethiopia, Mali, Mozambique, South Africa, Tanzania, Tunisia, Uganda, and Zambia.

^{xcix} Ghana, Malawi, Morocco, and Sierra Leone.

Table 35: Budget support effects on non-income poverty

Inputs		Total	BS Programme	Funds	Policy dialogue	Conditionality	TA / CD
Non-income Poverty	# Sources	33	32	9	2	0	0
	Total effect codings	72	58	12	2	0	0
	Positive effect	52	41	9	2	0	0
	No effect	19	16	3	0	0	0
	Negative effect	1	1	0	0	0	0
	Improving effect	0	0	0	0	0	0
	Deteriorating effect	0	0	0	0	0	0

district level.²²⁸

The evidence indicates that budget support did not manage to address the “missing middle”²²⁹ in service delivery, in terms of processes for management of frontline service providers, the actual delivery of services, human resources management, and the accountability for service provision.

Effects of the non-financial inputs policy dialogue and TA/CD are sufficiently covered, but coverage is nonetheless low and effects are inconsistent in their direction. Overall, positive effects are noted in a number of sources,²³⁰ particularly with regard to SBS. In other cases, however, authors find negative effects or no effects associated with the provision of non-financial inputs.²³¹ Shortcomings have been noted, particularly with respect to strengthening the administrative capacities for service delivery²³² and the development of frontline service delivery systems and staff.²³³

4.7.3 Non-income poverty

Coverage

For non-income poverty, effects of budget support are covered by 33 sources and discussed in 72 text sections (see Table 35). Generic effects of budget support are stated in 32 sources. Coverage for specific inputs is insufficient, with nine sources reporting effects of funding and two sources reporting effects of policy dialogue. Descriptions of the effects of conditionality and TA/CD on non-income poverty are not found in the course of the systematic coding process. Since there is insufficient evidence on the individual inputs’ effect on non-income

poverty, no generalizable conclusions on the effectiveness of specific budget support inputs, as formulated in the intervention logic, can be drawn. Knowledge of the effects of the specific budget support inputs that might inform future programmes or related instruments is thus lacking.

Established evidence

Despite a considerable number of sources (33) that address changes in non-income poverty, reliable evidence on the contribution of budget support to these changes is relatively scarce. It is important to note, however, that the majority of statements point in the same direction, suggesting improvements with regard to non-income poverty and drawing the distinct picture that there are no negative effects of budget support on non-income poverty.

While changes at impact level are described for most cases,²³⁴ attribution statements relating to the contribution of budget support to these trends is found only in a subset of sources (8 sources).²³⁵ This is because evaluations face problems in plausibly linking changes in non-income poverty to budget support programmes.^{236c} Although the evidence suggests a link, uncertainty prevails as to the contribution of budget support to positive trends in non-income poverty.

The evidence suggests that budget support does not reach its full potential, and explains the shortcoming with donors’ focus on quantitative targets to increase access to basic services,²³⁷ a lack of pro-poor policy orientation by the recipient government,²³⁸ dysfunctional accountability mechanisms,²³⁹

^c See also the IDD Evaluations, conducted in 2005 and published in 2006, focus on the output and outcome level with reference to the long term-nature of poverty impact (Lister, 2006).

and the weakness of budget support to affect local systems for service delivery.²⁴⁰

While the synthesis literature addresses the issues of gender equality²⁴¹ and general equality in access to public services,²⁴² sustainability of budget support funds²⁴³ and the effects of suspended budget support programmes on public service delivery,²⁴⁴ these issues were covered in fewer than 10 sources and are therefore not further discussed.



5.

CONCLUSION

This evaluation synthesis provides insights on what is known and what is not known about the effectiveness of budget support based on a much broader evidence-base than previous studies. By putting the main focus on the reliability of results, this evaluation synthesis clearly separates empirically established evidence from claims for which no sufficient empirical foundation exists and thus informs decision makers of both budget support programmes and related finance modalities. It answers the question of which effects of budget support are substantiated by reliable evidence and under which contextual conditions budget support generates results. It also shows for which outcomes evidence is not sufficient and thus highlights where further research is necessary to close remaining knowledge gaps.

5.1 Systematic approach

The evaluation synthesis focuses on the reliability of evidence by following a systematic approach, which consists of relevance criteria for search and inclusion of sources, a quality assessment, systematic coding and data analysis. The evaluation team systematically selected and reviewed 95 sources – evaluations as well as academic and grey literature. The evaluation synthesis thus offers broad coverage and the analysis of different perspectives on budget support effects. To control for qualitative differences in the evidence, a quality assessment based on a predefined set of quality criteria was conducted. The quality assessment showed the methodological quality of budget support reports and studies to be moderate but relatively homogeneous across sources, and all assessed sources were therefore deemed to be of sufficient methodological quality to be integrated as evidence in the evaluation synthesis. Despite the sufficient quality, the body of evidence was divided into best and second-best evidence on the basis of the quality assessment results. The threshold for best evidence was set at 50% of the maximum possible score. This facilitated a measure to examine the robustness of findings, a best-evidence robustness check. Findings are considered robust if they are covered by either best or by both best and second-best evidence. This was the case for all findings reported in this evaluation synthesis. Last, but not least, effects were only considered reliable and hence included as empirically established evidence if they were covered in 10

or more different sources.

5.2 Summary of findings

From the overall evidence on budget support effects analysed in this evaluation synthesis, six main findings emerge:

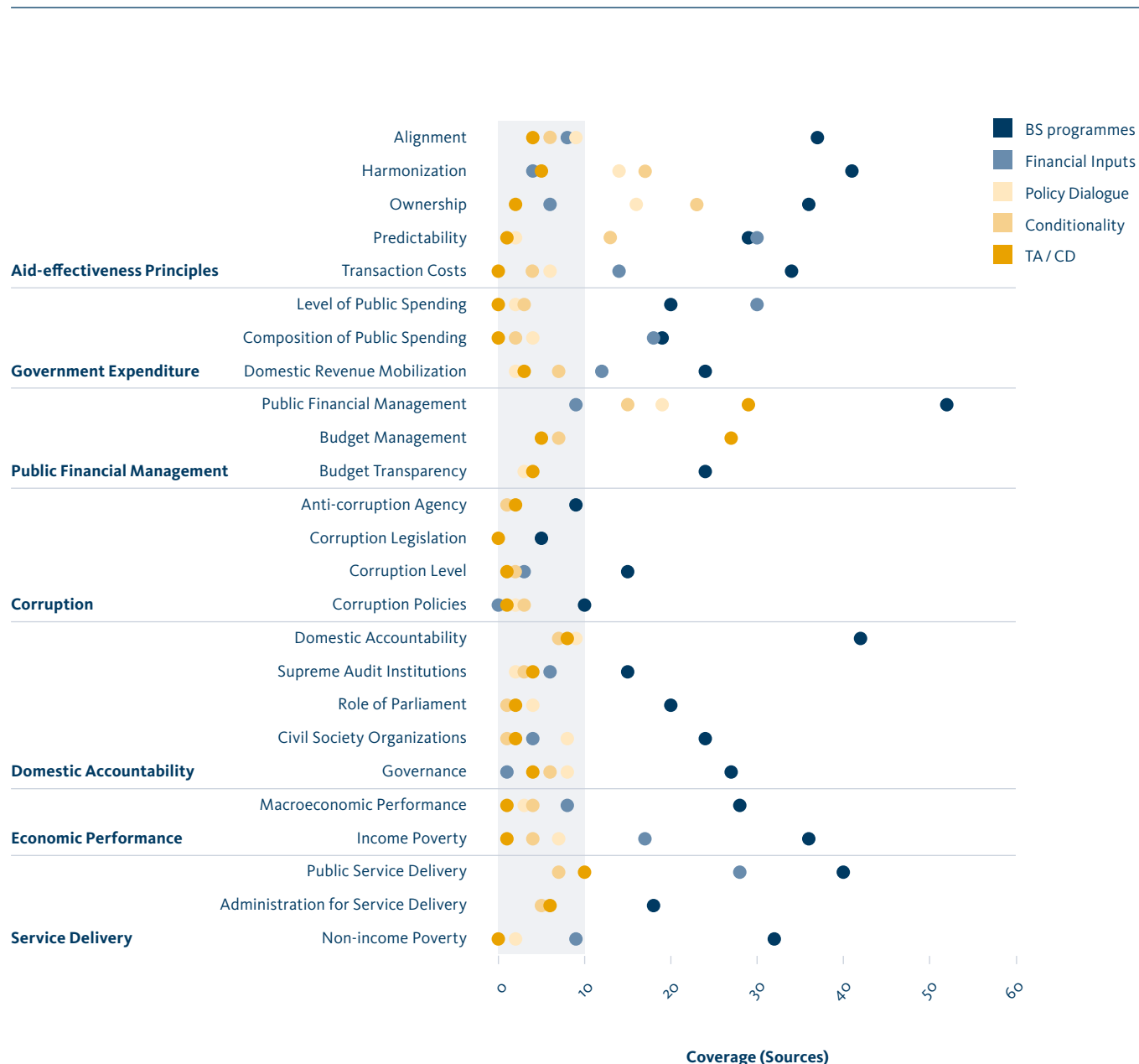
- Most expected effects of budget support are backed by sufficient evidence.
- Most proven effects of budget support are positive or at least in presence of certain context factors positive, such as a high number and quality of donors' common interests and the adherence of partner governments to the conditionalities.
- Observed effects are mostly attributed to the whole budget support programme. Information on the specific effects of individual inputs of budget support (financial input, conditionalities, policy dialogue and TA/CD) is rarely provided.
- A multitude of sources convincingly describes the attribution of observed changes to budget support on output and induced output level. On the outcome and especially the impact level, attribution of observed changes to budget support programmes is often not plausibly supported by evidence.
- Important and highly debated risks of budget support are inadequately researched by the covered sources, thus statements on risks are not possible.
- The strength of budget support effects depends strongly on context conditions, such as institutional capacity and political will for reform of the recipient government.

5.2.1 Coverage of budget support effects

A great majority of statements attribute observed changes to budget support programmes on a generic level. However, observed changes are not sufficiently attributed to one or a combination of the specific budget support inputs: funding, policy dialogue, conditionality and TA/CD.

The grey shaded area in Figure 6 indicates coverage below the benchmark of 10 reports for generic budget support effects (dark blue) or specific budget support inputs (blue and different shades of orange). While only a few generic budget

Figure 6: Coverage of budget support effects



support effects, namely on anti-corruption agencies and corruption legislation, are insufficiently covered by the evidence, most individual budget support inputs are insufficiently covered.

Thus, certain budget support effects that in some respect can be considered the centrepiece of the intervention logic of

budget support are barely covered in the body of evidence, such as the effect of non-financial inputs on the composition of public spending and domestic accountability. Exceptions with a sufficiently high coverage, for example, are effects of budget support funding on the level and composition of public spending, on DRM, public service delivery, and income poverty, and of non-financial inputs on PFM.

Although budget support's effects on corruption are widely discussed in the political sphere, the evidence insufficiently covers budget support effects on corruption in two out of four corruption codes. While the evidence sufficiently covers positive budget support effects on increases in the monitoring of corruption, and there is explicit evidence that budget support does not have a negative effect on corruption, it can generally not be assessed whether budget support contributes to a reduction in corruption or has no effect.

5.5.2 Established evidence

Figure 7 illustrates the direction of generic budget support effects on the outputs, outcomes, and impacts, as laid out in the intervention logic. The graph contrasts the assumptions on budget support effects formulated in the intervention logic with the evidence on budget support effects. While the graph indicates the direction of most budget support effects (positive, mixed, negative, no effect, implausible attribution, insufficient evidence), the effect on corruption can in part not be rated due to low coverage and is therefore not colour coded.

Positive effects

First, budget support has positive effects on the alignment of aid with national policies of partner countries, the ownership of partner governments in formulating and implementing own strategies for national development, and the harmonization of donor's aid policies. Budget support increases harmonization among donors under the condition that sector interests and priorities of donors align to a certain degree.

Second, budget support, especially budget support funding, increases the level of public spending. This increase is mostly apparent in the social sectors health and education.

Third, budget support programmes have a positive effect on the PFM of recipient countries. Improvements are, for example, reported in budget formulation and planning, in budgeting and accounting, and in the comprehensiveness and transparency of the budget. These improvements can specifically be attributed to the non-financial inputs policy dialogue and TA/CD.

Fourth, budget support generally contributes to improvements on the supply side of domestic accountability, mainly by channelling budget support funding through the national budget systems. Positive effects on domestic accountability include a strengthened role of supreme audit institutions, improvements in planning and managing national budget processes, and an increase in budget transparency.

Last, but not least, budget support has, as assumed in the intervention logic, a positive effect on people's access to public services. Funds from SBS and GBS are decisive to increase the quantity of service delivery, particularly in the education and health sectors.

Mixed effects

Most of the budget support effects show mixed results. The effectiveness of budget support in those cases is dependent on immanent factors, such as the programme cycle of budget support programmes or suspensions, and on context conditions. Mixed effects also refer to evaluation question 2: Under which contextual conditions does budget support generate results?

The evidence finds time-varying effects of budget support on transaction costs for partner governments. While transaction costs initially increase, they are subsequently reduced as processes become routinely established.

The absence of one suspected risk of budget support, the crowding-out of domestic revenue, is confirmed by the evidence. However, only a few sources identify positive effects of budget support programmes on revenue mobilization.

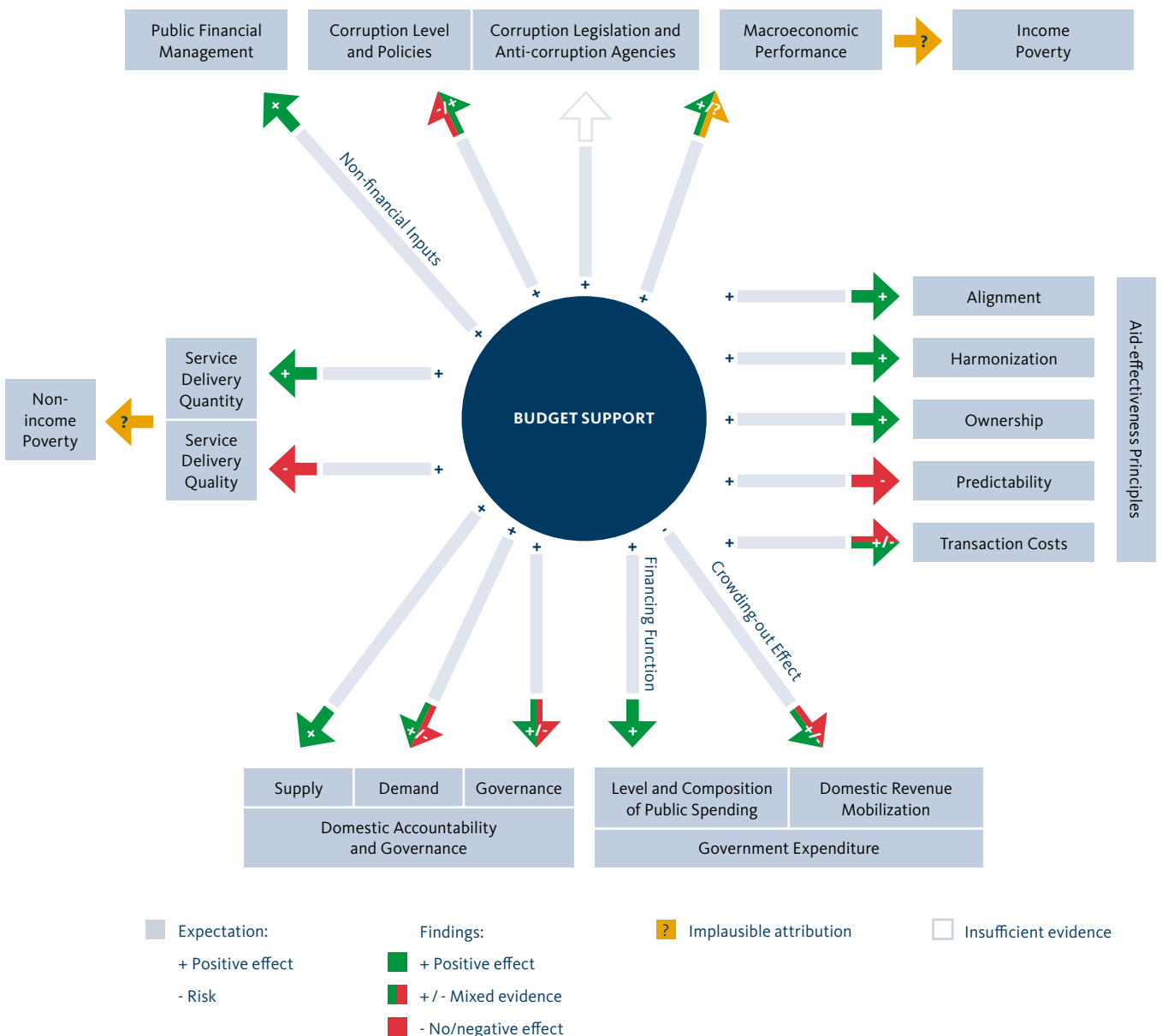
Whereas budget support increases domestic accountability on the supply side, effects are mixed on the demand side of domestic accountability and for democratic governance. Changes in parliament and CSO participation in the budget process are not sufficiently attributed to the presence of budget support programmes. Budget support programmes contribute to stronger intra-government accountability structures, whereas effects on decentralization are inconsistent across countries.

The evidence on macroeconomic performance suggests that budget support reinforces pre-existing macroeconomic stability. However, many sources do not attribute improvements in economic performance to budget support alone, but also, sometimes more prominently, to programmes of the IMF and

to the political commitment of the partner governments to achieve macroeconomic stability.

One factor limiting the effectiveness of budget support in a couple of outcomes is the suspension of budget support

Figure 7: Expectation and reality of budget support effects



disbursements. With suspensions of budget support, former gains in fiscal space and a government's room for manoeuvre diminished, enhanced by the low levels of DRM. Suspensions had mixed and even adverse effects on macroeconomic performance, and transaction costs started to increase. The extent to which budget support effects can be sustained after a majority of donors either suspend disbursements or stop their use of the aid modality will be analysed in detail in a second evaluation on budget support: the evaluation on the exit from budget support.

No effects

Budget support programmes, overall, do not improve predictability, due to late confirmations and disbursements of budget support tranches, and uncoordinated decisions by donors following breaches of conditionalities by partner governments.

While budget support programmes have a positive effect on people's access to public services, budget support is not found to increase the quality of delivered services.

Implausible attribution

The evidence on budget support effects on income poverty is mixed. Most sources find no or only slightly positive effects, but reliable evidence on the contribution of budget support to these changes is relatively scarce.

The evidence describes positive effects on non-income poverty in the recipient countries, but the effects are not attributed to budget support only.

5.3 Interpretation and recommendations

This evaluation synthesis presents the most comprehensive analysis to date of existing evidence on the effectiveness of budget support, a highly controversial modality of development cooperation. The findings systematically substantiate existing knowledge and generate new insights for decision makers in German and international development cooperation.

The study finds convincingly broad evidence that budget support is indeed an effective modality in promoting important development outcomes, such as improvements in public financial management and budget processes and improved provision of public goods and services. In view of these findings it appears worthwhile for donors – including those who have largely withdrawn from budget support – to re-assess the modality. In doing so, however, particular attention will need to be paid to the remaining evidence gaps identified in this evaluation synthesis. Specifically, there is a lack of knowledge with regard to the effectiveness of budget support at outcome and impact level due to unresolved methodological challenges and problems of attribution. Also, while the evidence proves effects of budget support programmes on a general level, little is known on the specific contribution of individual, particularly non-financial, budget support inputs. The same is true for hitherto largely neglected cross-cutting topics, such as budget support's effects on gender equality and distributional effects of growth.

More importantly, however, the evaluation synthesis also finds that – whereas positive effects of budget support are well covered in evaluations and other sources – the empirical evidence on potentially important fiduciary and other risks of budget support, such as corruption or adverse macro-economic effects (e.g. increases in public debt), remains surprisingly scarce. It thus seems fair to conclude that the collective body of evidence on budget support is somewhat skewed, not so much in terms of what it finds, but significantly so in where it looks: the body of evidence provides important findings of sufficiently high quality on the effects of budget support – as can be seen in the results of the quality assessment – but in doing so, it either does not refer to risks or adverse effects of budget support, or does so only

marginally. This is surprising insofar as the political debate on budget support in most donor countries tends to exhibit the opposite bias: whereas positive effects of budget support appear to have been largely ignored or dismissed as hypothetical and not proven in recent political discourse on the modality,²⁴⁵ the perceived risks of the aid modality received considerably more attention.²⁴⁶

The remaining gaps imply that, even from an extensive analysis of existing evidence such as this and despite strong evidence for positive effects of budget support in general, only limited conclusions can be drawn with regard to specific effects attributable to individual budget support inputs. Going hand in hand, there are only limited lessons to be learned for the design and implementation of aid modalities with similar features as budget support, such as results-based approaches. This also makes it difficult to predict the impact on development outcomes achieved through budget support of a widespread withdrawal of bilateral donors from the instrument.

Future empirical work therefore needs to closely analyse the effects and causal mechanisms of specific budget support inputs, as well as budget support effects on important cross-cutting issues. Also, future work should analyse in more depth in how far potential risks of budget support materialise in practice and how those risks can be mitigated.

To account for this need of more empirical work on the topic, DEval conducts a complementary evaluation of the exit from budget support, building on the findings of this evaluation synthesis and investigating the sustainability of budget support effects when donors suspend or exit from budget support. In conjunction with the findings presented here, this evaluation contributes to forming a more complete picture of the effectiveness of budget support, the consequences when donors exit from the aid modality, and the lessons for the design and implementation of future budget support programmes and related aid instruments.



6.

REFERENCES

Synthesized literature

Evaluation reports

EC/multi-donor evaluations

ADE (2015), *Evaluation conjointe de l'aide budgétaire au Burundi (2005-2013) Rapport Final*, EC; Belgium; France; Government of Burundi; ADE; ITAD; COWI, Brussels, <https://www.tresor.economie.gouv.fr/File/417799>.

Bartholomew, A. et al. (2006), *Evaluation of General Budget Support. Vietnam Country Report*, IDD; Mokoro; Ecorys; DRN; NCG, London; Glasgow, http://ec.europa.eu/europeaid/how/evaluation/evaluation_reports/reports/2006/705_vnm_en.pdf.

Batley, R. et al. (2006), *Evaluation of General Budget Support, Mozambique Country Report*, IDD; Mokoro; Ecorys; DRN; NCG, London; Glasgow, <http://www.oecd.org/countries/mozambique/43867765.pdf>.

Bogetic, Z. et al. (2015), *Joint Evaluation of Budget Support to Uganda. Final Report. Volume 1*, IEG; DFID; EC; Irish Aid; Government of Uganda, Brussels, https://ec.europa.eu/europeaid/sites/devco/files/strategic-evaluation-cooperation-ec-bs-uganda-1344-main-report-201505_en_o.pdf.

Booth, D. et al. (2005), *Joint Evaluation of General Budget Support Tanzania 1995–2004. Revised Final Report*, Daima Associates Limited; ODI, Dar es Salaam; London, <http://www.odi.org/sites/odi.org.uk/files/odi-assets/publications-opinion-files/3234.pdf>.

Caputo, E. et al. (2013), *Evaluation of Budget Support in South Africa. Final Report. Volume 1*, EC; Republic of South Africa; Particip; DRN; ecdpm; Ecorys; Mokoro, Brussels; Freiburg, http://ec.europa.eu/europeaid/how/evaluation/evaluation_reports/reports/2013/1322_vol1_en.pdf.

Claussen, J. et al. (2006), *Evaluation of General Budget Support, Malawi Country Report*, IDD, Birmingham, https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/67838/gbs-malawi.pdf.

de Kemp, A. et al. (2011a), *Between High Expectations and Reality: An Evaluation of Budget Support in Zambia - Synthesis Report*, BMZ; Ministry of Foreign Affairs; SIDA, Bonn; The Hague; Stockholm, <http://www.oecd.org/derec/sweden/zambia.pdf>.

Dijkstra, G. and A. Grigsby (2010), *Evaluation of General Budget Support to Nicaragua 2005–2008*, IOB, The Hague, <http://www.iob-evaluatie.nl/sites/iob-evaluatie.nl/files/329%20General%20Budget%20Support%20to%20Nicaragua%202005-08.pdf>.

DRN (2011), *Evaluation of European Commission Budget Support Operations in Tunisia between 1996 and 2008. Final Report. Volume 1*, EC; DRN; ADE; Particip; DIE; ODI; EURASIA; ICEI, Rome; Brussels, https://ec.europa.eu/europeaid/sites/devco/files/evaluation-budget-support-ec-tunisia-1286-main-report-201103_en.pdf.

Gosparini, P. et al. (2006), *Evaluation of General Budget Support. Nicaragua Country Report*, IDD; Mokoro; Ecorys; DRN; NCG, London; Glasgow, https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/67836/gbs-nicaragua.pdf.

Lanser, P. et al. (2006), *Evaluation of General Budget Support. Burkina Faso Country Report*, IDD; Mokoro; Ecorys; DRN; NCG, London; Glasgow, https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/67833/gbs-burkina-faso.pdf.

Lawson, A. et al. (2007), *Joint Evaluation of Multi-Donor Budget Support to Ghana. Final report. Volume 1*, ODI; Ghana Centre for Democratic Development, London, <http://www.odi.org/sites/odi.org.uk/files/odi-assets/publications-opinion-files/247.pdf>.

Lawson, A. et al. (2011), *Joint Evaluation of Budget Support Operations in Mali 2003–2009 Final Report Volume I*, EC; DGD; CIDA; ECO Consult; AGEG; APRI; Euronet; IRAM; NCG, Brussels, <https://www.oecd.org/countries/mali/48656665.pdf>.

Lawson, A. et al. (2013), *Joint Evaluation of Budget Support to Tanzania: lessons learned and recommendations for the future. Final Report: Volume 1*, DANIDA, EC; Irish Aid, Ministry of Foreign Affairs of the Netherlands; ADE; ITAD; COWI, Brussels, http://ec.europa.eu/europeaid/sites/devco/files/evaluation-budget-support-tanzania-1321-main-report-2013_en.pdf.

Lawson, A. et al. (2014), *Independent Evaluation of Budget Support in Mozambique. Final Report. Volume 1*, EC; Irish AID; AFC Ministry of Foreign Affairs of the Netherlands; BMZ; FORMIN; FARMESINA ADE; ITAD; COWI, Brussels, <http://www.oecd.org/derec/ec/Mz-BS-Eval-Final-Report-Vol-1.pdf>.

Lawson, A. et al. (2016), *Evaluation of Budget Support to Sierra Leone 2002–2015 Final Report Volume One*, DFID; EC; AfDB; World Bank; Ministry of Finance and Economic Development of Sierra Leone, London, https://ec.europa.eu/europeaid/sites/devco/files/evaluation-report-sierra-leone-volume-1-2016o62o_en.pdf.

Lister, S. et al. (2006), *Evaluation of General Budget Support. Uganda Country Report*, IDD; Mokoro; Ecorys; DRN; NCG, London; Glasgow, https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/67832/gbs-uganda.pdf.

Purcell, R. et al. (2006), *Evaluation of General Budget Support. Rwanda Country Report*, IDD; Mokoro; Ecorys; DRN; NCG, London; Glasgow, https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/67830/gbs-rwanda.pdf.

Thunnissen, K. and V. Morillon (2014), *Evaluation conjointe des opérations d'appui budgétaire au Maroc - Rapport final*, EC; ADE, Louvain-la-Neuve, https://ec.europa.eu/europeaid/sites/devco/files/strategic-evaluation-budget-support-morocco-1332-main-report-2014o9_fr_o.pdf.

Agency evaluations

AfDB (2011a), *Evaluation of Policy Based Operations in the African Development Bank, 1999–2009 Country Case Study: Sierra Leone*, <https://www.oecd.org/countries/sierraleone/49414435.pdf>.

[sierraleone/49414435.pdf](https://www.oecd.org/countries/sierraleone/49414435.pdf).

AfDB (2011b), *Evaluation of Policy Based Operations in the African Development Bank, 1999–2009 Country Case Study: Tanzania*, http://www.afdb.org/fileadmin/uploads/afdb/Documents/Evaluation-Reports-_Shared-With-OPEV_/Tanzania%20case%20study%20FINAL.pdf.

IDB (2015), *Design and Use of the Policy-Based Loans at the IDB*, <https://publications.iadb.org/handle/11319/7513?locale-attribute=en>.

World Bank (2008), *Project Performance Assessment Report Ethiopia*, [http://lnweb90.worldbank.org/oed/oeddoelib.nsf/24cc3bb1f94ae11c85256808006a0046/a35a3f335eed78a285257ba300755579/\\$FILE/PPAR43336.pdf](http://lnweb90.worldbank.org/oed/oeddoelib.nsf/24cc3bb1f94ae11c85256808006a0046/a35a3f335eed78a285257ba300755579/$FILE/PPAR43336.pdf).

World Bank (2009a), *Project Performance Assessment Report Burkina Faso*, <http://documents.worldbank.org/..../475480PPARoPo71CoDisclosedo41241091.pdf>.

World Bank (2009b), *Project Performance Assessment Report Mozambique*, <http://documents.worldbank.org/curated/en/611351468280492987/pdf/489200PPARoPo41CoDisclosedo7171091.pdf>.

World Bank (2009c), *Project Performance Assessment Report Uganda*, <http://documents.worldbank.org/curated/en/332711468175737633/pdf/489420PPARo1mu1o1OfficialoUseoOnly1.pdf>.

World Bank (2013a), *Project Performance Assessment Report Dominican Republic*, <http://documents.worldbank.org/curated/en/100661468248707343/pdf/787870PPARoP1100Box377365BooPUBLICo.pdf>.

World Bank (2013b), *Project Performance Assessment Report Tanzania*, <http://documents.worldbank.org/curated/en/542381468133184866/pdf/781880PPARoTanooBox377382BooPUBLICo.pdf>.

World Bank (2014), *Project Performance Assessment Report Santa Lucia*, <http://documents.worldbank.org/curated/>

en/311331468303048351/pdf/882400PPARoP110Codisclosedo60110140.pdf.

World Bank (2015a), *Project Performance Assessment Report Uganda*, <http://documents.worldbank.org/curated/en/942871468188377758/pdf/96202-PPAR-P090881-P090219-P101231-P101232-P097325-Box385353B-OUO-9-IDA-SecM2015-0150.pdf>.

World Bank (2015b), *Project Performance Assessment Report Vietnam*, <http://documents.worldbank.org/curated/en/721331467986245693/pdf/96203-PPAR-P101724-P105287-P111164-P111182-P111183-P104097-P107062-P117610-P117723-P120946-box393169B-OUO-9-IDA-R2015-0210.pdf>.

Academic literature

Alavuotunki, K. (2015), “General budget support, health expenditures, and neonatal mortality rate. A synthetic control approach”, *Working Paper*, No. 2015/108, UNU-WIDER, <https://www.wider.unu.edu/sites/default/files/wp2015-108.pdf>.

Bernardi, M. et al. (2015), *EU State Building Contracts. Early Lessons from the EU’s New Budget Support Instrument for Fragile States*, ODI, <https://www.odi.org/sites/odi.org.uk/files/odi-assets/publications-opinion-files/9441.pdf>.

Bird, N. and C. Avoka (2007), *Budget Support, Aid Instruments and the Environment: The Country Context: Ghana Country Case Study*, ODI; CDD-Ghana, <http://www.odi.org/sites/odi.org.uk/files/odi-assets/publications-opinion-files/4665.pdf>.

Brown, A. et al. (2000), *Aid Transaction Costs in Viet Nam*, ODI, London, http://www.swisstph.ch/fileadmin/user_upload/Pdfs/swap/swap157.pdf.

Celasum, O. and J. Walliser (2005), *Predictability of Budget Aid: Experiences in Eight African Countries*, IMF; World Bank, <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.518.7004&rep=rep1&type=pdf>.

de Renzio, P. et al. (2010), *Evaluation of Donor Support to Public Financial Management (PFM) Reform in Developing Countries: Analytical Study of Quantitative Cross-country Evidence*, ODI, <https://www.odi.org/sites/odi.org.uk/files/odi-assets/publications-opinion-files/6373.pdf>.

Del Biondo, K. and J. Orbie (2014), “The European Commission’s implementation of budget support and the governance incentive tranche in Ethiopia: Democracy promoter or developmental donor?”, *Third World Quarterly*, Vol. 35, No. 3: 411–427.

Dijkstra, G. et al. (2012), “Budget support: Conditional results. Review of an instrument (2000–2011)”, *IOB Evaluation*, No. 369, IOB, The Hague, http://www.oecd.org/derec/netherlands/IOB_BS.pdf.

Dijkstra, G. (2013), “Governance or poverty reduction? Assessing budget support in Nicaragua”, *Journal of Development Studies*, Vol. 49, No. 1: 110–124.

Eifert, B. and A. Gelb (2005), “Improving the dynamics of aid: Towards more predictable budget support”, *Policy Research Working Paper*, No. 3732, World Bank, https://papers.ssrn.com/sol3/papers.cfm?abstract_id=822844

European Court of Auditors (2010), *The Commission’s Management of General Budget Support in ACP, Latin American and Asian countries*, European Court of Auditors, Luxembourg, http://www.eca.europa.eu/Lists/ECADocuments/SR10_11/SR10_11_EN.PDF.

Faust, J. et al. (2012), “Financing poverty alleviation vs. promoting democracy? Multi-donor budget support in Zambia”, *Democratization*, Vol. 19, No. 3: 438–464.

Faust, J. and S. Koch (2014), “Foreign aid and the domestic politics of European budget support”, *Discussion Paper*, No. 21/2014, Deutsches Institut für Entwicklungspolitik (DIE), Bonn, https://www.die-gdi.de/uploads/media/DP_21.2014.pdf.

- Fernandes Antunes, A. et al. (2010)**, “General budget support: Has it benefited the health sector?”, *World Health Report (2010) Background Paper*, No. 14, WHO, <http://www.scopus.com/inward/record.url?eid=2-s2.0-84887233764&partnerID=40&md5=6e435b5c82701246382c7a958do6e3e2>.
- Fisher, J. (2011)**, *Budget Support and Good Governance in Uganda*, IDD, <http://eadi.org/gc2011/fisher-312.pdf>.
- Hayman, R. (2011)**, “Budget support and democracy: A twist in the conditionality tale”, *Third World Quarterly*, Vol. 32, No. 4: 673–688.
- Hedger, E. et al. (2010)**, *Sector Budget Support in Practice. Case Study Education Sector in Uganda*, ODI; Mokoro, London, <http://www.odi.org/sites/odi.org.uk/files/odi-assets/publications-opinion-files/6077.pdf>.
- Inberg, L. and N. Holvoet (2013)**, “Multiple pathways to gender-sensitive budget support in the education sector: Analysing the effectiveness of sex-disaggregated indicators in performance assessment frameworks and gender working groups in (education) budget support to Sub-Saharan Africa countries”, *Working Paper*, No. 2013/105, UNU-WIDER, <https://www.econstor.eu/dspace/bitstream/10419/93681/1/770116191.pdf>.
- Knoll, M. (2011)**, “Foreign aid and revenue response: An examination of joint general budget support”, *Working Paper*, No. 2011/23, School of Business & Economics Discussion Paper: Economics, Free University Berlin, School of Business & Economics, Berlin, <http://www.econstor.eu/bitstream/10419/54729/1/680587365.pdf>.
- Koch, S. and P. Morazan (2010)**, *Monitoring Budget Support in Developing Countries: A Comparative Analysis of National Control Mechanisms over Budget Support in Developing Countries*, European Parliament, http://www.europarl.europa.eu/RegData/etudes/etudes/join/2010/410210/EXPO-DEVE_ET%282010%29410210_EN.pdf.
- Kremser, R. (2011)**, *Einflussnahme der Geber im Rahmen von Budgethilfe-Dialogmechanismen auf die Entwicklung von Politiken und Strategien: Fallstudie Sektor Landwirtschaft in Mosambik (ProAgri)*, Universität Wien, Wien.
- Lawson, A. (2007)**, *DFID Budget Support to Sierra Leone, 2004–2007: Achievements & Lessons for the Future*, Report to the UK Department for International Development, ODI, <http://www.odi.org/sites/odi.org.uk/files/odi-assets/publications-opinion-files/249.pdf>.
- Lawson, A. and S. Bouaré (2008)**, *Budget Support, Aid Instruments and the Environment: The Country Context: Mali Country Case Study*, ODI, <https://www.cbd.int/financial/expenditure/mali-budgetsupport.pdf>.
- Leiderer, S. et al. (2007)**, “Public Financial Management for PRSP implementation in Malawi: Formal and informal PFM institutions in a decentralising system”, *Studies*, No. 28, Deutsches Institut für Entwicklungspolitik (DIE), Bonn, https://www.die-gdi.de/uploads/media/Studies_28.pdf.
- Leiderer, S. et al. (2012)**, “Efficiency of service provision at local government level in Zambia in health, education and roads. Implications for decentralization and the effectiveness of budget support”, *Studies*, No. 71, Deutsches Institut für Entwicklungspolitik (DIE), Bonn, http://saipar.org:8080/eprc/bitstream/handle/123456789/305/GIZ_Efficiency%20of%20Local%20Service%20Provision%20in%20Zambia%27s%20Health%2c%20Education%20and%20Road%20Sectors_2012.pdf?sequence=1&isAllowed=y.
- Leiderer, S. and J. Faust (2012)**, “Evaluation of budget support in Zambia. Implementation, direct effects and political economy”, *Studies*, No. 68, Deutsches Institut für Entwicklungspolitik (DIE), Bonn, http://edoc.vifapol.de/opus/volltexte/2013/4406/pdf/Studies_68.pdf.
- Lister, S. (2006)**, *Evaluation of General Budget Support: Synthesis Report*, IDD; Mokoro; Ecorys; DRN; NCG, https://www.bmz.de/en/publications/type_of_publication/evaluation/international_joint_evaluations/Synthese.pdf.

- Molenaers, N. et al. (2010)**, “Budget support and policy/political dialogue: Donor practices in handling (political) crises”, *Discussion Paper*, No. 2010/06, IOB; University of Antwerp, <http://s3.amazonaws.com/academia.edu.documents/31029207/00256641.pdf?AWSAccessKeyId=AKIAJ56TQJRTWSMTNPEA&Expires=1441178851&Signature=bOnmb4TUYgzlDWYWLw%2BErWNLD%2B4%3D&response-content-disposition=inline>.
- Molenaers, N. (2012)**, “The Great Divide? Donor perceptions of budget support, eligibility and policy dialogue”, *Third World Quarterly*, Vol. 33, No. 5: 791–806.
- Molenaers, N. et al. (2015)**, “What determines the suspension of budget support?”, *World Development*, Vol. 75: 62–73.
- Mosley, P. and A. Suleiman (2005)**, “Budget support, conditionality and poverty”, *Working Paper*, No. 2005012, Sheffield Economic Research Paper Series, University of Sheffield, Department of Economics, Sheffield, <http://eprints.whiterose.ac.uk/9922/1/SERP2005012.pdf>.
- Nordtveit, I. (2012)**, *Herd Behavior or Increased Efficiency: What Determines the Allocation of General Budget Support?*, University of Bergen, Department of Economic, http://economics.gu.se/digitalAssets/1373/1373871_nordtveit.pdf.
- Nordtveit, I. (2014)**, “Does better governance and commitment to development attract general budget support?”, *Working Paper*, No. 2/14, University of Bergen, Department of Economic, Bergen, <http://www.uib.no/sites/w3.uib.no/files/attachments/02-2014.pdf>.
- ODI and Mokoro (2009)**, *Sector Budget Support in Practice: Case Study Education Sector in Rwanda*, ODI; Mokoro, <https://www.odi.org/sites/odi.org.uk/files/odi-assets/publications-opinion-files/5580.pdf>.
- Schmidt, P. (2006)**, “Budget support in the EC’s development cooperation”, *Studies*, No. 20, Deutsches Institut für Entwicklungspolitik (DIE), Bonn, <http://dspace.africaportal.org/jspui/bitstream/123456789/26066/1/Budget%20support%20in%20the%20ECs%20development%20cooperation.pdf?1>.
- Smith, G. (2009)**, *Sector Budget Support in Practice: Desk Study Health Sector in Tanzania*, ODI; Mokoro, <http://www.odi.org/sites/odi.org.uk/files/odi-assets/publications-opinion-files/6403.pdf>.
- Steffensen, J. (2010)**, *Sector Budget Support in Practice: Desk Study Local Government Sector in Uganda*, ODI; Mokoro, <http://www.odi.org/sites/odi.org.uk/files/odi-assets/publications-opinion-files/6407.pdf>.
- Swedlund, H.J. (2013)**, “From donorship to ownership? Budget support and donor influence in Rwanda and Tanzania”, *Public Administration and Development*, Vol. 33, No. 5: 357–370.
- Swedlund, H.J. (2014)**, “Assessing the promises of budget support: Case study evidence from Rwanda”, *Problems, Promises, and Paradoxes of Aid: Africa’s Experience*, Cambridge Scholars Publishing, Newcastle upon Tyne: 238–264.
- Tidemand, P. (2009)**, *Sector Budget Support in Practice: Desk Study Local Government Sector in Tanzania*, ODI; Mokoro, <http://www.odi.org/sites/odi.org.uk/files/odi-assets/publications-opinion-files/5583.pdf>.
- Williamson, T. and C. Dom (2010)**, *Sector Budget Support in Practice: Synthesis Report*, ODI; Mokoro, <http://www.odi.org/sites/odi.org.uk/files/odi-assets/publications-opinion-files/5806.pdf>.
- Woll, B. (2008)**, “Donor harmonisation and government ownership: Multi-donor budget support in Ghana”, *European Journal of Development Research*, Vol. 20, No. 1: 74–87.

Grey literature

Castel-Branco, C.N. et al. (2008), *Mozambique Programme Aid Partners Performance Review 2007*, IESE, http://www.iese.ac.mz/lib/publication/outras/PAPs_PAF_2007.pdf.

Chigunta, F. and N. Matshalaga (2010), *Evaluation of the Implementation of the Paris Declaration in Zambia*, <http://search.oecd.org/dac/evaluation/dcdndep/44549817.pdf>.

Dom, C. and A. Gordon (2011), "Budget support in fragile situations", *Discussion Paper*, Oxfam; Mokoro, <https://www.oxfam.org/sites/www.oxfam.org/files/dp-budget-support-fragile-situations-111111-en.pdf>.

EU (2014), *EU Budget Support to Eastern Partnership Countries: Cases of Moldova, Georgia and Ukraine*, OSF; REC Moldova; Sweden, http://eap-csf.eu/assets/files/WG1_EU%20Budget%20support_last_en.pdf.

Frantz, B. (2004), *General Budget Support in Tanzania: A Snapshot of Its Effectiveness*, USAID, http://www.swisstph.ch/fileadmin/user_upload/Pdfs/swap/swap384.pdf.

Gerster, R. and M. Chikwekwe (2007), *Poverty Reduction Budget Support (PRBS) in Zambia Joint Annual Review 2007: Learning Assessment*, Report to the Government of the Republic of Zambia and the PRBS Cooperating Partners, Gerster Consulting, http://www.gersterconsulting.ch/docs/LA_Zambia.pdf.

Gerster, R. and R.G. Mutakyahwa (2006), *Annual Review 2006 of General Budget Support in Tanzania: Learning Assessment*, Report to the Government of Tanzania and The Development Partners' Group, Gerster Consulting, http://www.researchgate.net/profile/Richard_Gerster/publication/237478316_Annual_Review_2006_of_General_Budget_Support_in_Tanzania_Learning_Assessment/links/54e252eaocf2edaeao92bd6a.pdf.

Gerster, R. and K. Sawadogo (2003), *Independent Evaluation of SDC's Bilateral Engagement in the Poverty Reduction Strategy Paper (PRSP) Process. Part 2: Case Studies: Burkina Faso*, Gerster Consulting; Development Initiatives, http://www.gersterconsulting.ch/docs/report_burkina_faso.pdf.

Harding, A. and R. Gerster (2004), *Learning Assessment of Joint Review 2004*, Report to the Government of Mozambique and Programme Aid Partners, Gerster Consulting, http://www.gersterconsulting.ch/docs/la_report.pdf.

Jimat Consult (2011), *Phase II Evaluation of the Implementation of the Paris Declaration in Uganda*, Jimat Development Consultants; Government of the Republic of Uganda, <https://www.oecd.org/countries/uganda/47673315.pdf>.

Lawson, A. et al. (2003a), *General Budget Support Evaluability Study Phase 1: Synthesis Report, Volume 1*, DFID Evaluation Report, OPM; ODI, <http://www.odi.org/sites/odi.org.uk/files/odi-assets/publications-opinion-files/4835.pdf>.

Lawson, A. et al. (2003b), *General Budget Support Evaluability Study: Country Case Studies: Mozambique @ Andhra Pradesh, Volume II*, DFID Evaluation Report, OPM; ODI, <http://www.odi.org/sites/odi.org.uk/files/odi-assets/publications-opinion-files/4837.pdf>.

Lawson, A. et al. (2005), *Learning from Experience with Performance Assessment Frameworks for General Budget Support*, OPM; Gerster Consulting; SECO, http://gersterconsulting.ch/docs/Synthesis_PAF_Report.pdf.

Lieberson, J. et al. (2004), "General budget support and sector program assistance: Malawi country case study", *PPC Evaluation Working Paper*, No. 19, USAID, http://www.swisstph.ch/fileadmin/user_upload/Pdfs/swap/swap454.pdf.

Loquai, C. and H. Klavert (2011), "Supporting domestic accountability in the context of budget support", *Briefing Note*, No. 28, ecdpm, <http://ecdpm.org/wp-content/uploads/2013/11/BN-28-Domestic-Accountability-Budget-Support-German-Development-Malawi-2011.pdf>.

Molt, P. and G. Wahlers (2008), *Budgethilfe als Instrument der Entwicklungszusammenarbeit: Wirkungen und Herausforderungen in Ghana, Mosambik, den palästinensischen Autonomiegebieten und Vietnam*, Konrad-Adenauer-Stiftung, http://www.kas.de/wf/doc/kas_15210-544-1-30.pdf.

Ofori-Atta, G. (2012), “General budget support, facilitating aid effectiveness through harmonization and reduced transaction costs: A case study of Ghana’s multi-donor budget support (MDBS)”, *Research Paper*, Erasmus University, International Institute of Social Studies, The Hague, <http://thesis.eur.nl/pub/13061/>.

Orlowski, D. (2013), *Impact of Variable Tranches on Policy and Reforms. Background Note in the Context of Preparation of the Danish Budget Support Contract with Mozambique for 2015-2017*, Royal Danish Embassy Mozambique, Dudenrode.

Puonti, P. (2010), *Interplay of Aid and Fiscal Policy: Does Budget Support Induce Greater Fiscal Discipline?*, Master’s Thesis, Aalto University, Department of Economics, Espoo, https://aaltodoc.aalto.fi/bitstream/handle/123456789/531/hse_thesis_12399.pdf?sequence=1.

Rao, S. and H. Marquette (2012), “Corruption indicators in Performance Assessment Frameworks for budget support”, *U4 Issue*, No. 1/2012, U4; CMI, <http://www.cmi.no/publications/file/4400-corruption-indicators-in-performance-assessment.pdf>.

Reinsberg, B. (2012), *Rewarding Democratization? The Dynamics of General Budget Support in the Wake of Democratic Transitions*, University of Zurich, Institute for Political Science, Zurich, http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2182260.

Works cited

Acharya, A. et al. (2006), “Proliferation and fragmentation. Transaction costs and the value of aid”, *Journal of Development Studies*, Vol. 42, No. 1: 1–21.

Aldasoro, I. et al. (2009), “Less aid proliferation and more donor coordination? The wide gap between words and deeds”, *Working Paper*, No. 1516, Kiel Institute for the World Economy, Kiel, https://www.files.ethz.ch/isn/100018/kwp_1516.pdf.

Alton-Lee, A. (2004), *Using Best Evidence Syntheses to Assist in Making a Bigger Difference for Diverse Learners*, Ministry of Education, New Zealand, Wellington, <http://www.educationalleaders.govt.nz/Pedagogy-and-assessment/Building-effective-learning-environments/Using-Best-Evidence-Syntheses-for-diverse-learners>.

Bigsten, A. (2006), “Donor coordination and the uses of aid”, *Working Papers in Economics*, Göteborg University, Göteborg, <https://gupea.ub.gu.se/bitstream/2077/2723/1/gunwpeo196.pdf>.

Birdsall, N. et al. (2010), *Cash on Delivery: A New Approach to Foreign Aid*, CGD Books.

Campbell Collaboration (2015), *Campbell Collaboration Systematic Reviews: Policies and Guidelines*, The Campbell Collaboration, https://www.campbellcollaboration.org/images/C2_Policies_and_Guidelines_Doc_Version_1_1.pdf.

Caputo, E. et al. (2011), *Application of New Approach to the Evaluation of Budget Support Operations: Findings from Mali, Zambia and Tunisia: Synthesis of Main Results, EC*, <https://www.oecd.org/countries/zambia/49716695.pdf>.

de Kemp, A. et al. (2011b), “Too much, too quickly: The woes of budget support”, *The Broker*: 16.

DeGEval (2008), *Standards für Evaluation*, Mainz, 4th ed., http://www.degeval.de/fileadmin/user_upload/Sonstiges/STANDARDS_2008-12.pdf.

- DFID (2004)**, “Poverty reduction budget support”, *DFID Policy Paper*, DFID, <http://webarchive.nationalarchives.gov.uk/+/> <http://www.dfid.gov.uk/pubs/files/prbspaper.pdf>.
- DFID (2014)**, *Sharpening Incentives to Perform: DFID’s Strategy for Payment by Results*, DFID, https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/323868/Sharpening_incentives_to_perform_DFIDs_Strategy_on_Payment_by_Results.pdf.
- Dijkstra, G. and A. de Kemp (2015)**, “Challenges in evaluating budget support and how to solve them”, *Evaluation*, Vol. 21, No. 1: 83–98.
- Easterly, W. (2005)**, “What did structural adjustment adjust? The association of policies and growth with repeated IMF and World Bank adjustment loans”, *Journal of Development Economics*, Vol. 76: 1–22.
- EC (2008a)**, *Budget Support. A Question of Mutual Trust*, EC, Belgium, http://www.pedz.uni-mannheim.de/daten/edz-k/dev/o8/LM_budget_support_en.pdf.
- EC (2008b)**, “Methodology for evaluations of budget support operations at country level”, *Issue Paper*, DRN; ADE; ECO-Consulting; NCG; Ecorys, Rome; Brussels, http://ec.europa.eu/europeaid/how/evaluation/evaluation_reports/reports/2008/1258_isspap_en.pdf.
- EC (2012)**, “Budget support guidelines: Programming, design and management – A modern approach to budget support”, *Working Document, Tools and Methods Series*, EC, Brussels, https://ec.europa.eu/europeaid/sites/devco/files/methodology-budget-support-guidelines-201209_en_2.pdf.
- EC (2015)**, *Budget Support – Financial Implementation, Risk Assessment, and Selected Poverty, Macroeconomic and Fiscal Results*, EC, Brussels.
- European Court of Auditors (2014)**, “EuropeAid’s evaluation and results-oriented monitoring systems”, *Special Report*, No. 18, European Court of Auditors, Luxembourg, http://www.eca.europa.eu/Lists/ECADocuments/SR14_18/SR14_18_EN.pdf.
- European Parliament et al. (2005)**, *The European Consensus on Development. Joint statement by the Council and the Representatives of the Governments of the Member States meeting with the Council, the European Parliament and the Commission on the European Union Development Policy*, Brussels, http://ec.europa.eu/development/body/development_policy_statement/docs/edp_declaration_signed_20_12_2005_en.pdf#zoom=125.
- Faust, J. et al. (2011)**, “Multi-donor budget support: Only halfway to effective coordination”, *Briefing Paper*, No. 8/2011, Deutsches Institut für Entwicklungspolitik (DIE), Bonn, https://www.die-gdi.de/uploads/media/BP_8.2011.pdf.
- Faust, J. et al. (2012)**, *The Future of EU Budget Support: Political Conditions, Differentiation and Coordination*, ODI, London, http://docs.google.com/viewerng/viewer?url=http://ecdpm.org/wp-content/uploads/EU-Budget-Support-Political-Conditions-Differentiation-Coordination.pdf&hl=en_US&safe=strict.
- Faust, J. and S. Koch (2014)**, “Foreign Aid and the Domestic Politics of European Budget Support”, *Discussion Paper*, No. 21/2014, Deutsches Institut für Entwicklungspolitik (DIE), Bonn, https://www.die-gdi.de/uploads/media/DP_21.2014.pdf.
- Frot, E. and J. Santiso (2009)**, “Crushed aid: Fragmentation in sectoral aid”, *Working Paper*, No. 6/2009, Stockholm Institute of Transition Economics, Stockholm, <https://core.ac.uk/download/pdf/6385525.pdf?repositoryId=153>.
- Glock, B. and K.-M. Karliczek (2014)**, *Meta-Evaluation, Evaluationssynthese und Entwicklung von Bewertungskriterien, Camino – Werkstatt für Fortbildung, Praxisbegleitung und Forschung im sozialen Bereich gGmbH, Berlin*, http://www.jugendgewaltpraevention.de/sites/default/files/Metaevaluation_und_Evaluationskriterien.pdf.
- Halonen-Akatwijuka, M. (2004)**, “Coordination Failure in Foreign Aid”, Policy Research Working Paper, No. 3223, World Bank, Washington D.C., <http://documents.worldbank.org/curated/en/831971468762915890/pdf/wps3223AID.pdf>.

- Handley, G. (2009)**, “Ask the Right Questions”, *D+C, Development and Cooperation*, No. 4, <http://www.dandc.eu/en/article/right-questions-concerning-budget-support-review-essay>.
- Janus. H. (2014)**, “Real Innovation or Second-Best Solution? First Experiences from Results-Based Aid for Fiscal Decentralisation in Ghana and Tanzania”, *Discussion Paper*, No. 3/2014, Deutsches Institut für Entwicklungspolitik (DIE), Bonn, https://www.die-gdi.de/uploads/media/DP_3.2014.pdf.
- Janus, H. and S. Klingebiel (2014)**, “Results-based aid: Potential and limits of an innovative modality in development cooperation”, *International Development Policy*, No. 5.2.
- Kayizzi-Mugerwa, S. (1998)**, “Africa and the donor community: From conditionality to partnership”, *Journal of International Development*, Vol. 10, No. 2: 219–225.
- Klingebiel, S. (2012)**, “Results-based aid (RBA): New aid approaches, limitations and the application to promote good governance”, *Discussion Paper*, No. 14/2012, Deutsches Institut für Entwicklungspolitik (DIE), Bonn, https://www.die-gdi.de/uploads/media/DP_14.2012.pdf.
- Knack, S. and A. Rahman (2007)**, “Donor fragmentation and bureaucratic quality in aid recipients”, *Journal of Development Economics*, Vol. 83, No. 1: 176–197.
- Knoll, M. (2008)**, “Budget support: a reformed approach or old wine in new skins?”, *Discussion Paper*, No. 190, United Nations Conference on Trade and Development, Geneva, http://unctad.org/en/docs/osgdp20085_en.pdf. Koch. S. et al. (2016), “The rise and demise of European budget support: Political economy of collective European Union donor action”, *Development Policy Review*: 1–19.
- Koch. S. (2015)**, “A typology of political conditionality beyond aid: Conceptual horizons based on lessons from the European Union”, *World Development*, Vol. 75/2015 (Special Issue: Political Conditionality and EU Foreign Aid), pp. 98–108.
- Koeberle, S. et al. (2006)**, *Budget Support as More Effective Aid? Recent Experiences and Emerging Lessons*, World Bank Publications, Washington D.C.
- Krisch, F. et al. (2015)**, *Accompanying Measures to General Budget Support in Sub-Saharan Africa*, Deutsches Evaluierungsinstitut der Entwicklungszusammenarbeit, Bonn, p. 112.
- Lawson, A. (2014)**, *Synthesis of Budget Support Evaluations: Analysis of the Findings, Conclusions and Recommendations of Seven Country Evaluations of Budget Support*, ADE, Fiscus, EC, https://ec.europa.eu/europeaid/sites/devco/files/strategic-evaluation-synthesis-ec-budget-support-1335-main-report-201411_en.pdf.
- Lawson, A. and D. Booth (2004)**, *Evaluation Framework for General Budget Support*, Overseas Development Institute, London, https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/67937/genbudgetsup.pdf.
- Leiderer, S. (2010)**, “Budget support as an aid instrument: neither pandemonium nor panacea”, *Briefing Paper*, No. 9/2010, Deutsches Institut für Entwicklungspolitik (DIE), Bonn, http://www.die-gdi.de/uploads/media/BP_9.2010.pdf.
- Leiderer, S. (2012)**, “Fungibility and the choice of aid modalities – The red herring revisited”, *Working Paper*, No. 2012/68, UNU-WIDER, Helsinki, <https://www.wider.unu.edu/sites/default/files/wp2012-068.pdf>.
- Littell, J.H. and J. Corcoran (2010)**, “Systematic reviews”, in *The Handbook of Social Work Research Methods*, SAGE Publications, Inc., Thousand Oaks, 2nd ed.: 313.
- Littell, J.H. and B.R. Maynard (2014)**, “The science of research synthesis: Limiting bias and error in reviews”, *Paper*, No. 59, Graduate School of Social Work and Social Research Faculty Research and Scholarship, http://repository.brynmawr.edu/cgi/viewcontent.cgi?article=1058&context=gsswsr_pubs.

- Nilsson, M. (2004)**, “Effects of budget support – A discussion of early evidence”, *UTV Working Paper*, SIDA, Stockholm, <https://www.oecd.org/dac/evaluation/dcdndep/35031496.pdf>.
- OECD DAC (1991)**, *Principles for Evaluation of Development Assistance*, OECD/DAC, Paris, <https://www.oecd.org/development/evaluation/2755284.pdf>.
- OECD DAC (2005)**, *The Paris Declaration on Aid Effectiveness and the Accra Agenda for Action*, OECD/DAC, Paris, <http://www.oecd.org/dac/effectiveness/34428351.pdf>.
- OECD DAC (2006)**, *Harmonising Donor Practices for Effective Aid Delivery. Volume 2: Budget Support, Sector Wide Approaches and Capacity Development in Public Financial Management*, OECD/DAC Guidelines and Reference Series, Paris: 1–81.
- OECD DAC (2008)**, *2008 Survey on Monitoring the Paris Declaration: Making Aid More Effective by 2010*, OECD/DAC, Paris, <http://www.oecd.org/dac/effectiveness/2008surveyonmonitoringtheparisdeclaration.htm>.
- OECD DAC (2009)**, *Development Co-operation Report 2009*, OECD/DAC, Paris, <http://www.oecd.org/dac/42193704.pdf>.
- OECD DAC (2012)**, *Evaluating Budget Support. Methodological Approach*, OECD/DAC, https://www.oecd.org/dac/evaluation/dcdndep/Methodological%20approach%20BS%20evaluations%20Sept%202012%20_with%20cover%20Thi.pdf.
- Popay, J. et al. (2006)**, *Guidance on the Conduct of Narrative Synthesis in Systematic Reviews. A Product from the ESRC Methods Programme*, Lancaster University.
- Rønsholt, F.E. (2014)**, “Review of budget support evaluations”, *Evaluation Study*, No. 2014/1, DANIDA, Copenhagen, <https://www.oecd.org/derec/denmark/Review-of-Budget-Support-Evaluation.pdf>.
- Saini, M. and A. Shlonsky (2012)**, *Systematic Synthesis of Qualitative Research*, Oxford University Press, New York, <http://www.oxfordscholarship.com/view/10.1093/acprof:oso/9780195387216.001.0001/acprof-9780195387216-chapter-005>.
- Schmitt, J. and D. Beach (2015)**, “The contribution of process tracing to theory-based evaluations of complex aid instruments”, *Evaluation*, Vol. 21, No. 4: 429–447.
- Stern, E. et al. (2012)**, “Broadening the range of designs and methods for impact evaluations”, *DFID Working Paper*, No. 38, DFID, London, <https://assets.publishing.service.gov.uk/media/57a08a6740f0b6497400059e/DFIDWorkingPaper38.pdf>.
- Stufflebeam, D.L. (1999)**, *Program Evaluation Models Metaevaluation Checklist (Based on The Program Evaluation Standards)*, https://www.wmich.edu/sites/default/files/attachments/u350/2014/eval_model_metaeval.pdf.
- Tavakoli, H. and G. Smith (2013)**, “Back under the microscope: Insights from evidence on budget support”, *Development Policy Review*, Vol. 31, No. 1: 59–74.
- UNCTAD (2014)**, *World Investment Report 2014: Investing in the SDGs: An Action Plan*, UNCTAD, Geneva, http://unctad.org/en/PublicationsLibrary/wir2014_en.pdf.



7.

ANNEX

7.1 Three-step approach

The Three-Step approach²⁴⁷ is the most used approach in evaluations using the Comprehensive Evaluation Framework. The approach acknowledges that effects of budget support can be traced up to the level of induced outputs of the CEF, while the attribution of effects on outcomes and impact level to budget support is very difficult. It also takes into account that the induced outputs (level 3) are also influenced by other factors than the budget support programmes and that outcomes and impact (level 4 and 5) are part of the effects of the partner governments' policy performance, depending on the national poverty reduction strategies, the institutional quality and the political will for reforms.

The three steps of the approach are:

1. Aid effects evaluation

The first step is an assessment of the inputs, direct outputs and induced outputs (level 1, 2, and 3 in the CEF) of the budget support programmes and an evaluation of the causal link between these levels, including the external factors influencing the effects in level 3.

2. Policy impact evaluation

The second step is an analysis of the outcomes and impact (level 4 and 5 in the CEF) of the national policies that were supported by budget support programmes. This analysis includes identifying factors (possibly direct and induced outputs) influencing these outcomes and impact.

3. Qualitative contribution analysis

The third step combines the results of steps 1 and 2, to identify the effect of budget support on the outcomes and impact (level 4 and 5). Because the link between budget support and outcomes and impact is not straightforward, this analysis is done qualitatively, based on logical reasoning.

The three steps do not have to be followed in chronological order. However, steps 1 and 2 always have to precede step 3.

7.2 Quality criteria evaluation reports (including scaling system)

Category	Quality criterion	Rating
Intervention Logic and Evaluation Questions	QC 1: Formulation of evaluation questions (EQ)	0 = no EQ formulated 1 = one general EQ formulated 2 = multiple EQs formulated 3 = more than 5 EQs formulated covering at least three levels of the intervention logic
	QC 2: The evaluation answers the formulated EQs	0 = no link between conclusions and EQ 1 = some EQs are answered 2 = all EQs are partly answered 3 = all EQs are fully answered
Context	QC3: The evaluation describes the economic and political context in the case country	0 = not at all 1 = described only to a limited extent 2 = socio-economic context described 3 = comprehensive context analysis
	QC4: The evaluation describes the context of aid relations in the case country	0 = not at all 1 = described only to a limited extent 2 = aid relation described 3 = comprehensive context analysis
Methodology	QC5: The evaluation makes use of an intervention logic (IL)	0 = no IL applied 1 = generic IL applied 2 = country specific IL applied 3 = country specific IL reconstructed with stakeholders
	QC 6: The evaluation acknowledges problems of attribution and describes how it addresses them in the methodology section	0 = no discussion 1 = limited discussion 2 = discussion and solutions presented 3 = convincing discussion and adequate solutions * give reasons
	QC 7: The evaluation reflects on the influence of (changing) context on programme outcomes in the methodology section	0 = no reflection 1 = monochronic and general 2 = monochronic and specific for EQs OR diachronic and general 3 = diachronic and specific for EQs
	QC 8: The evaluation describes how it triangulates methods and sources in the methodology section	0 = no triangulation described 1 = triangulation of either methods OR sources described 2 = triangulation of methods AND sources described
	QC 9: The report describes general limitations of the evaluation in the methodology section	0 = no discussion 1 = some limitations listed 2 = some limitations discussed 3 = comprehensive discussion of limitations

Category	Quality criterion	Rating
Data Collection and Analysis	QC 10: The evaluation provides information on the collected data (figures, interview partners, document references)	0 = no information 1 = some information 2 = comprehensive information 3 = 2+reflects on the condition of data collection
	QC 11: The scope of the empirical inquiry (stakeholder groups)	0 = no interviews conducted 1 = interviews cover only one perspective (donor or recipient) 2 = interviews cover both perspectives 3 = interviews cover multiple factions from both perspectives
	QC 12: The scope of the empirical inquiry (number of interviews)	0 = no interviews conducted 1 = 1 to 10 2 = 11 to 30 3 = 31 and more
	QC 13: The evaluation provides information on the procedures for data analysis	0 = no information 1 = some information 2 = comprehensive information 3 = 2+evaluation documents how information from each procedure was scored, analysed, and interpreted
Plausible Conclusion	QC 14: The evaluation presents its conclusions with reference to methodological limitations	0 = limitations not mentioned at all 1 = limitations mentioned in the conclusions 2 = specific limitations discussed for individual conclusions
	QC 15: The evaluation makes reference to the information that supports each conclusion	0 = no reference made 1 = sporadic references but not systematic for all conclusion 2 = systematic reference 3 = fully transparent (e.g. using an evidence grid)
	QC 16: The evaluation identifies and discusses the programme's unintended effects	0 = no information 1 = unintended effects identified 2 = unintended effects identified and discussed 3 = unintended effects part of the evaluation design (in EQ) and systematically reported
	QC 17: The evaluation elaborates plausible alternative explanations of the findings.	0 = no alternative explanations elaborated 1 = some are elaborated 2 = alternative explanations for individual findings elaborated
	QC 18: The evaluation justifies why rival explanations were rejected	0 = rejection of rival explanations not justified 1 = rejection of rival explanations justified

7.3 Overview on budget support programmes

Country	Evaluation Period	Size of Donor Group	Joint Assistance Framework	Net ODA as % of GNI (period average)	BS as % of government expenditure (period average)
Burkina Faso	1994–2004	9	yes	14,9063346	6,91%
Burundi	2005–2013	8	no	29,4475301	17,79%
Ghana	2003–2006	12	yes	11,6232235	16,39%
Malawi	1994–2004	8	yes	22,9248008	3,54%
Mali	2003–2009	10	yes	11,6185121	8,18%
Morocco	2005–2012	6	no	1,38735183	0,24%
Mozambique	2005–2012	19	no	18,3217492	22,77%
Mozambique	1994–2004	17	no	28,209749	13,19%
Nicaragua	2005–2008	9	yes	11,073732	10,45%
Nicaragua	1994–2004	14	yes	15,8793815	8,65%
Rwanda	1994–2004	9	no	30,4292521	23,51%
Sierra Leone	2002–2015	6	no	19,4207975 (without 2015)	27,57% (without 2015)
South Africa	2000–2011	11	no	0,33869829	0,02%
Tanzania	2006–2012	14	yes	9,5033332	16,07%
Tanzania	1995–2004	14	in development	12,8363009	12,73%
Tunisia	1996–2008	1/3	joint programs	1,06853849	0,23%
Uganda	2004–2013	12	yes	11,2638873	12,42%
Uganda	1994–2004	12/20	planned	13,6741075	12,99%
Vietnam	1994–2004	17	no	4,31553652	5,56%
Zambia	2005–2010	9	yes	9,34195719	50,31% (only 2010 available)

8.

ENDNOTES

1. Acharya et al., 2006; Aldasoro et al., 2009; Bigsten, 2006; Frot/Santiso, 2009; Halonen-Akatwijuka, 2004; Knack/Rahman, 2007; OECD DAC, 2009.
2. Easterly, 2005/Kayizzi-Mugerwa, 1998.
3. OECD DAC, 2005.
4. Koeberle et al., 2006.
5. Klingebiel 2012; Janus, 2014.
6. Acharya et al., 2006; Aldasoro et al., 2009; Bigsten, 2006; Frot/Santiso, 2009; Halonen-Akatwijuka, 2004; Knack/Rahman, 2007; OECD DAC, 2009.
7. Easterly, 2005/Kayizzi-Mugerwa, 1998.
8. OECD DAC, 2005.
9. Koeberle et al., 2006.
10. Klingebiel, 2012; Janus, 2014.
11. Koeberle et al., 2006; Leiderer, 2010; OECD DAC, 2005.
12. Aldasoro et al., 2009; Bigsten, 2006; Frot/Santiso, 2009; Halonen-Akatwijuka, 2004; Knack/Rahman, 2007; OECD DAC, 2009.
13. Easterly, 2005; Kayizzi-Mugerwa, 1998.
14. OECD DAC, 2005.
15. OECD DAC, 2008: 148.
16. Dijkstra/de Kemp, 2015.
17. OECD DAC, 2006: 26.
18. Williamson/Dom, 2010.
19. Faust et al., 2011.
20. DFID, 2004; EC, 2008a.
21. European Parliament et al., 2005.
22. de Kemp et al., 2011b; Leiderer, 2010.
23. Dijkstra/de Kemp, 2015; Krisch et al., 2015.
24. Koeberle et al., 2006.
25. Faust et al., 2011.
26. Faust/Koch, 2014.
27. EC, 2012.
28. EC, 2015.
29. Klingebiel, 2012.
30. Birdsall et al., 2010; Leiderer, 2012.
31. UNCTAD, 2014.
32. Klingebiel, 2012; Janus, 2014.
33. Dijkstra/de Kemp 2015.
34. Lawson et al., 2003a.
35. Bartholomew et al., 2006; Batley et al., 2006; Claussen et al., 2006; Gosparini et al., 2006; Lanser et al., 2006; Lister et al., 2006; Purcell et al., 2006.
36. Lister, 2006.
37. Leiderer, 2012.
38. EC, 2008b.
39. de Kemp et al., 2011a; DRN, 2011; Lawson et al., 2011.
40. Bogetic et al., 2015; Lawson et al., 2013; Lawson et al., 2014.
41. Dijkstra et al., 2012; Lawson, 2014; Rønsholt, 2014; Tavakoli/Smith, 2013.
42. Faust et al., 2012; Krisch et al., 2015; OECD DAC, 2012.
43. OECD DAC, 2006.
44. EC, 2008b.
45. de Kemp et al., 2011a; DRN, 2011; Lawson et al., 2011.
46. EC, 2008b; Lawson et al., 2003a; OECD DAC, 2012.
47. OECD DAC, 2012.
48. OECD DAC, 2012: 11.
49. Schmitt/Beach, 2015.
50. European Court of Auditors, 2014; Schmitt/Beach, 2015; Stern et al., 2012.
51. Dijkstra et al., 2012; Lawson, 2014; Rønsholt, 2014; Tavakoli/Smith, 2013.
52. Saini/Shlonsky, 2012: 6.
53. Littell/Corcoran, 2010 cited in Saini/Shlonsky, 2012: 6.
54. Saini/Shlonsky 2012.
55. Saini/Shlonsky 2012.
56. Popay et al., 2006; Campbell Collaboration, 2015.
57. Alton-Lee, 2004; DeGEval, 2008; Glock/Karliczek, 2014; Stufflebeam, 1999.
58. Koeberle et al., 2006; Lawson et al., 2003a.
59. AfDB, 2011b; Batley et al., 2006; Bernardi, 2015; Bogetic et al., 2015; Booth et al., 2005; Caputo et al., 2013; Chigunta/Matshalaga, 2010; Claussen et al., 2006; de Kemp et al., 2011a; Dijkstra/Grigsby, 2010; DRN, 2011; Gerster/Chikwekwe, 2007; Gosparini et al., 2006; Koch/Morazan, 2010; Lanser et al., 2006; Lawson et al., 2007; Lister et al., 2006; Loquai/Klavert, 2011; ODI/Mokoro, 2009; Ofori/Atta, 2012; Purcell et al., 2006; Swedlund, 2014; World Bank, 2008; World Bank, 2013a; World Bank, 2014; World Bank, 2015b.
60. AfDB, 2011b; Batley et al., 2006; Booth et al., 2005; Caputo et al., 2013; Dijkstra/Grigsby, 2010; Gerster/Chikwekwe, 2007; Lanser et al., 2006; Lawson et al., 2007; Lister et al., 2006; Loquai/Klavert, 2011; ODI/Mokoro, 2009; World Bank, 2008.

61. Bernardi, 2015; Bogetic et al., 2015; Chigunta/Matshalaga, 2010; Claussen et al., 2006; de Kemp et al., 2011a; Lister, 2006; Lister et al., 2006; ODI/Mokoro, 2009; Ofori/Atta, 2012; World Bank, 2013a; World Bank, 2014; World Bank, 2015b.
62. AfDB, 2011b; Chigunta/Matshalaga, 2010; DRN, 2011; Gerster/Chikwekwe, 2007; ODI/Mokoro, 2009; Swedlund, 2014.
63. Bogetic et al., 2015; Claussen et al., 2006; de Kemp et al., 2011a.
64. Claussen et al., 2006; Dijkstra et al., 2012; Koch/Morazan, 2010; Lanser et al., 2006; ODI/Mokoro, 2009; Purcell et al., 2006; Swedlund, 2014.
65. de Kemp et al., 2011a.
66. AfDB, 2011b; Batley et al., 2006; Bogetic et al., 2015; Booth et al., 2005; Claussen et al., 2006; de Kemp et al., 2011a; Dijkstra, 2013; Dijkstra/Grigsby, 2010, DRN, 2011, Gerster/Chikwekwe, 2007; Gosparini et al., 2006; Hedger et al., 2010; Koch/Morazan, 2010; Lanser et al., 2006; Lawson, 2007; Lawson et al., 2003a; Lawson et al., 2007; Lawson et al., 2011; Lister, 2006; Purcell et al., 2006; Swedlund, 2014.
67. Bogetic et al., 2015; Chigunta/Matshalaga, 2010; de Kemp et al., 2011a; Dijkstra/Grigsby, 2010; Dom/Gordon, 2011; DRN, 2011; Faust et al., 2012; Gerster/Mutakyahwa, 2006; Lawson et al., 2005; Lawson et al., 2011; Ofori/Atta, 2012; World Bank, 2008; World Bank, 2015a.
68. Bogetic et al., 2015; de Kemp et al., 2011a; Dom/Gordon, 2011; DRN, 2011; Faust et al., 2012.
69. Bogetic et al., 2015; Dijkstra/Grigsby, 2010; Lawson et al., 2011; World Bank, 2008.
70. Bogetic et al., 2015; Claussen et al., 2006; de Kemp et al., 2011a; Dijkstra/Grigsby, 2010; Lister et al., 2006; Purcell et al., 2006.
71. AfDB, 2011b; Batley et al., 2006; Bogetic et al., 2015; Claussen et al., 2006; Faust et al., 2012; Faust/Koch, 2014; Harding/Gerster, 2004; Knoll, 2011; Lanser et al., 2006; Lawson et al., 2005; Lawson et al., 2007; Lawson et al., 2013; ODI/Mokoro, 2009; Schmidt, 2006.
72. Batley et al., 2006; Claussen et al., 2006; DRN, 2011; Koeberle et al., 2006; Lawson et al., 2011; Leiderer/Faust, 2012; ODI/Mokoro, 2009; Ofori/Atta, 2012; World Bank, 2015a.
73. Lawson et al., 2016.
74. AfDB, 2011b; Batley et al., 2006; Claussen et al., 2006; Harding/Gerster, 2004; Knoll, 2011; Lanser et al., 2006; Lawson et al., 2005; Lawson et al., 2007; Lawson et al., 2013; ODI/Mokoro, 2009.
75. Bogetic et al., 2015; Claussen et al., 2006; Faust et al., 2012; Faust/Koch, 2014; Schmidt, 2006.
76. Tavakoli/Smith, 2013.
77. Booth et al., 2005; Caputo et al., 2013; de Kemp et al., 2011a; Dijkstra/Grigsby, 2010; Dijkstra et al., 2012; Gerster/Chikwekwe, 2007; Harding/Gerster, 2004; Lanser et al., 2006; Lawson et al., 2003a; Lawson et al., 2005; Lister, 2006; Molt/Wahlers, 2008; ODI/Mokoro, 2009; Ofori/Atta, 2012; Swedlund, 2013; Swedlund, 2014; Williamson/Dom, 2010; Woll, 2008; World Bank, 2013a.
78. Claussen et al., 2006; Lawson et al., 2003b; Molt/Wahlers, 2008; World Bank, 2008; World Bank, 2014.
79. Lister, 2006; Lister et al., 2006; Steffensen, 2010.
80. Dijkstra/Grigsby, 2010; Lawson et al., 2003a.
81. Bogetic et al., 2015; Chigunta/Matshalaga, 2010; Lawson et al., 2013; World Bank, 2015b.
82. ADE, 2015; Batley et al., 2006; Claussen et al., 2006; de Kemp et al., 2011a; Harding/Gerster, 2004; Gerster/Chikwekwe, 2007; Kremser, 2011; Lawson et al., 2007; Molenaers et al., 2010; Molt/Wahlers, 2008; Ofori/Atta, 2012; Purcell et al., 2006.
83. Batley et al., 2006; Harding/Gerster, 2004; Molt/Wahlers, 2008; Ofori/Atta, 2012.
84. Batley et al., 2006; Harding/Gerster, 2004; Gerster/Chikwekwe, 2007; Lawson et al., 2007; Molt/Wahlers, 2008.
85. Chigunta/Matshalaga, 2010; de Kemp et al., 2011a; Gosparini et al., 2006.
86. de Kemp et al., 2011a; Swedlund, 2013; Swedlund, 2014.
87. de Kemp et al., 2011a; Lawson et al., 2003a; Molt/Wahlers, 2008.
88. Batley et al., 2006; Caputo et al., 2013; Celasum/Walliser, 2005; Dijkstra/Grigsby, 2010; Frantz, 2004; Lanser et al., 2006; Lawson et al., 2014; Molt/Wahlers, 2008; Ofori/Atta, 2012; Purcell et al., 2006; World Bank, 2008; World Bank, 2009a.
89. Batley et al., 2006; Dijkstra, 2013; Dijkstra/Grigsby, 2010; Frantz, 2004; Gosparini et al., 2006; Harding/Gerster, 2004; Lawson et al., 2013; Lawson et al., 2014; Molt/Wahlers, 2005; ODI/Mokoro, 2009; Purcell et al., 2006; Swedlund, 2014.

90. AfDB, 2011b; Bartholomew et al., 2006; Bogetic et al., 2015; Claussen et al., 2006; de Kemp et al., 2011a; Dom/Gordon, 2011; Gosparini et al., 2006; Harding/Gerster, 2004; Jimat Consult, 2011; Lawson, 2007; Lawson et al., 2003a; Lawson et al., 2016; Leiderer et al., 2007; Leiderer/Faust, 2012; Lieberson et al., 2004; Swedlund, 2014.
91. Batley et al., 2006; Caputo et al., 2013; Celasum/Walliser, 2005; Dijkstra/Grigsby, 2010; Frantz, 2004; Lanser et al., 2006; Lawson et al., 2014; Molt/Wahlers, 2008; Ofori/Atta, 2012; Purcell et al., 2006; World Bank, 2008.
92. Bartholomew et al., 2006; Batley et al., 2006; Chigunta/Matshalaga, 2010.
93. Batley et al., 2006; Dijkstra, 2013; Dijkstra/Grigsby, 2010; Gosparini et al., 2006; Harding/Gerster, 2004; Lawson et al., 2014; Molt/Wahlers, 2005; ODI/Mokoro, 2009; Purcell et al., 2006; Swedlund, 2014.
94. Claussen et al., 2006; Dijkstra, 2013; Dijkstra/Grigsby, 2010; Gosparini et al., 2006; Rao/Marquette, 2012.
95. Claussen et al., 2006; Lawson et al., 2003a; Leiderer et al., 2007.
96. AfDB, 2011a; Claussen et al., 2006; Lawson et al., 2003a; Leiderer et al., 2007; Lister, 2006; Molt/Wahlers, 2008; Orłowski, 2013; Purcell et al., 2006; Rao/Marquette, 2012; Smith, 2009.
97. Chigunta/Matshalaga, 2010; Gerster/Chikwekwe, 2007; Lawson, 2007.
98. Bartholomew et al., 2006; Batley et al., 2006; Claussen et al., 2006; de Kemp et al., 2011a; Dijkstra/Grigsby, 2010; Dijkstra et al., 2012; Dom/Gordon, 2011.
99. AfDB, 2011b; Bartholomew et al., 2006; Bogetic et al., 2015; Claussen et al., 2006; de Kemp et al., 2011a; Dom/Gordon, 2011; Gosparini et al., 2006; Harding/Gerster, 2004; Jimat Consult, 2011; Lawson, 2007; Lawson et al., 2003b; Lawson et al., 2016; Leiderer et al., 2007; Leiderer/Faust, 2012; Lieberson et al., 2004; Swedlund, 2014.
100. Bartholomew et al., 2006; Batley et al., 2006; Claussen et al., 2006; Dijkstra/Grigsby, 2010; Frantz, 2004; Gerster/Mutakyahwa, 2006; Gosparini et al., 2006; Jimat Consult, 2011; Lawson et al., 2003a; Lawson et al., 2005; Lawson et al., 2007; Lawson et al., 2011; Lawson et al., 2014; Lawson et al., 2016; Leiderer/Faust, 2012; Lister et al., 2006; Molt/Wahlers, 2008; Ofori/Atta, 2012; Purcell et al., 2006; Tidemand, 2009; Williamson/Dom, 2010; World Bank, 2008.
101. Batley et al., 2006; Dijkstra/Grigsby, 2010; Frantz, 2004; Gerster/Mutakyahwa, 2006; Gosparini et al., 2006; Lawson et al., 2003b; Lawson et al., 2005; Lawson et al., 2007; Lister et al., 2006; Molt/Wahlers, 2008; Ofori/Atta, 2012; Tidemand, 2009; Williamson/Dom, 2010.
102. Batley et al., 2006; Frantz, 2004; Gosparini et al., 2006; Jimat Consult, 2011; Lawson et al., 2003a; Lawson et al., 2013; Lawson et al., 2014; Lawson et al., 2016; Leiderer/Faust, 2012; Molt/Wahlers, 2008; Purcell et al., 2006.
103. Tavakoli/Smith, 2013.
104. Gosparini et al., 2006; Lawson et al., 2013; Lawson et al., 2014; Lawson et al., 2016; Leiderer/Faust, 2012; Molt/Wahlers, 2008.
105. Lawson et al., 2014; Purcell et al., 2006.
106. Dijkstra, 2013; Dijkstra et al., 2012; Dijkstra/Grigsby, 2010; Lawson et al., 2013; Lister et al., 2006.
107. Claussen et al., 2006.
108. de Kemp et al., 2011a.
109. World Bank, 2009c.
110. Tavakoli/Smith, 2013.
111. Leiderer/Faust, 2012.
112. ADE, 2015; Batley et al., 2006; Frantz, 2004; Lawson, 2007; Lawson et al., 2007; Lawson et al., 2013; Lister et al., 2006; Molt/Wahlers, 2008; World Bank, 2015a.
113. ADE, 2015; Bogetic et al., 2015; Claussen et al., 2006.
114. Lanser et al., 2006; Lawson, 2007; Lawson et al., 2016; Ofori-Atta, 2012; Purcell et al., 2006; World Bank, 2009a.
115. Lawson et al., 2011; Lister, 2006; Lister et al., 2006; Ofori-Atta, 2012; Williamson/Dom, 2010; World Bank, 2009a; World Bank, 2009c.
116. Lawson 2011; Lister, 2006; Lister et al., 2006.
117. Batley et al., 2006; Booth et al., 2005; Caputo et al., 2013; de Kemp et al., 2011a; Lanser et al., 2006; Lawson et al., 2006; Lawson et al., 2012; Lawson et al., 2013; Lawson et al., 2016; Lister et al., 2006; Ofori/Atta, 2012; Thunnissen/Morillon, 2014.
118. ADE, 2015; Bartholomew et al., 2006; Batley et al., 2006; Bogetic et al., 2015; Booth et al., 2005; Chigunta/Matshalaga, 2010; Claussen et al., 2006; de Kemp et al., 2011a; Dijkstra/Grigsby, 2010; Dijkstra et al., 2012; DRN, 2011; Faust et al., 2012; Frantz, 2004; Lawson, 2007; Lawson et al., 2003a; Lawson et al., 2007; Lawson et al., 2013; Lawson et al., 2014; Lister et al., 2006; Thunnissen/Morillon, 2014; World Bank, 2008; World Bank, 2009a; World Bank, 2009c; World Bank, 2013a; World Bank, 2015a; World Bank, 2015b; Williamson/Dom, 2010.

119. ADE, 2015; Lawson et al., 2007; Lister, 2006.
120. Booth et al, 2005; Caputo et al., 2013; Lawson et al., 2006; Lawson et al., 2012; Lawson et al., 2013; Lawson et al., 2016; Lister et al., 2006; Ofori/Atta, 2012; Thunnissen/Morillon, 2014.
121. ADE, 2015; Caputo et al., 2013; de Kemp et al., 2011a; Lawson et al., 2012; Lawson et al., 2016.
122. Bogetic et al., 2015; Claussen et al., 2006.
123. ADE, 2015; AfDB, 2011a; Bartholomew et al., 2006; Batley et al., 2006; Bogetic et al., 2015; Chigunta/Matshalaga, 2010; Dijkstra, 2013; Dijkstra et al., 2012; Dijkstra/Grigsby, 2010; Knoll, 2011; Lawson et al., 2011; Leiderer/Faust, 2012; Lister, 2006; Lister et al., 2006; Orłowski, 2013; World Bank, 2009a; World Bank, 2013b; World Bank, 2015a.
124. ADE, 2015; DRN, 2011; Knoll, 2011; Lanser et al., 2006; Lawson et al., 2013; Lawson et al., 2014.
125. Frantz, 2004; Lanser et al., 2006; Lawson et al., 2016; World Bank, 2015a.
126. Bogetic et al., 2015; Lawson et al., 2016; World Bank, 2013b.
127. Bogetic et al., 2015; Claussen et al., 2006; Leiderer et al., 2007.
128. Bogetic et al., 2015; Purcell et al., 2006; World Bank, 2015a.
129. Dijkstra et al., 2012; Knoll, 2011.
130. Batley et al, 2006; de Kemp et al., 2011a; Dijkstra, 2013; Dijkstra et al., 2012; Knoll, 2011; Lawson, 2007; Lawson et al., 2014; Leiderer/Faust, 2012; Lister, 2006; Lister et al., 2006.
131. Dijkstra et al., 2012; Nilsson, 2004; Tavakoli/Smith, 2013.
132. Batley et al., 2006; Bogetic et al., 2015; Booth et al. 2005; de Kemp et al., 2011a; Dijkstra, 2013; Dijkstra/Grigsby, 2010; DRN, 2011; Gosparini et al., 2006; Lanser et al., 2006; Lawson, 2007; Lawson et al., 2007; Lawson et al., 2011; Lawson et al., 2013; Lawson et al., 2014; Lawson et al., 2016; Lister et al, 2006; ODI/Mokoro, 2009; Purcell et al., 2006; Thunnissen/Morillon, 2014; World Bank, 2015a; World Bank, 2015b.
133. Batley et al., 2006; Bogetic et al., 2015; Booth et al., 2005; European Court of Auditors, 2010; Fernandes Antunes et al., 2010; Frantz, 2004; Harding/Gerster, 2004; Knoll, 2011; Lawson et al., 2016; Lister, 2006; Lister et al., 2006; Loquai/Klavert, 2011; ODI/Mokoro, 2009; Rønsholt, 2014; Steffensen, 2010; Swedlund, 2014.
134. AfDB, 2011a; Bogetic et al., 2015; Claussen et al., 2006; Dijkstra, 2013; Lawson, 2016; Leiderer/Faust, 2012.
135. Bogetic et al., 2015; Claussen et al., 2006; de Kemp et al., 2011a; Gosparini et al., 2006; Lawson, 2007; Lawson et al., 2014.
136. Bartholomew et al., 2006; Claussen et al. 2006; Lawson, 2007; Lister et al., 2006; Thunnissen/Morillon, 2014.
137. Bartholomew et al., 2006; de Kemp et al., 2011a; Lawson, 2007; Loquai/Klavert, 2011.
138. AfDB, 2011b; Hedger et al., 2010; Lawson et al., 2007; Lawson et al., 2016; Smith, 2009; World Bank, 2015a.
139. Bogetic et al., 2015; Gerster/Chikwekwe, 2007; Lawson, 2007; Lawson et al., 2014; Lister et al., 2006; ODI/Mokoro, 2009; Tidemand, 2009.
140. Claussen et al., 2006; de Kemp et al., 2011a; Gosparini et al., 2006; Lanser et al., 2006; Lawson et al., 2011; Lawson et al., 2013; Lawson et al., 2016.
141. Bogetic et al., 2015.
142. Lawson et al., 2007; Leiderer/Faust, 2012; World Bank, 2008; World Bank, 2015a.
143. IOB, 2012; Lawson et al., 2007; World Bank 2008.
144. Koeberle et al., 2006; OECD DAC, 2012.
145. Bogetic et al., 2015; Knoll, 2011; Koch/Morazan, 2010; Lawson et al., 2013.
146. AfDB, 2011a; Batley et al., 2006; Bogetic et al., 2015; Booth et al., 2005; de Kemp et al., 2011a; Dijkstra, 2013; DRN 2011; European Court of Auditors, 2010; Frantz, 2004; Knoll, 2011; Koch/Morazan, 2010; Lawson, 2007; Lawson et al 2011; Lawson et al., 2013; Lawson et al., 2014; Lawson et al., 2016; Purcell et al., 2006; Swedlund, 2014; Tidemand, 2009.
147. de Kemp et al., 2011a; Purcell et al., 2006.
148. AfDB, 2011a; Batley et al., 2006; Bogetic et al., 2015; de Kemp et al., 2011a; Dijkstra, 2013; DRN, 2011; European Court of Auditors, 2010; Lawson et al., 2014; Lawson et al., 2016; Knoll, 2011.
149. AfDB, 2011b; de Kemp et al., 2011a; DRN, 2011; Lawson et al., 2014; Lawson et al., 2016.
150. de Kemp et al., 2011a; EU, 2014; Lawson et al., 2014; Lawson et al., 2016.
151. Batley et al., 2006; de Kemp et al., 2011a.
152. Dijkstra et al., 2012; Lawson, 2007; Lawson et al., 2013; Leiderer/Faust, 2012; Purcell et al. 2006, Williamson/Dom, 2010.
153. Lawson, 2007.

154. Bartholomew et al., 2006; Batley et al., 2006; Bernardi et al., 2015; Booth et al., 2005; Caputo et al., 2013; Dijkstra/Grigsby, 2010; DRN, 2011; Lanser et al., 2006; Lawson, 2007; Lawson et al., 2007; Lawson et al., 2011; Lawson et al., 2013; Lawson et al., 2014; Lister et al., 2006; Purcell et al., 2006; Thunnissen/Morillon, 2014; Williamson/Dom, 2010.
155. Dijkstra et al., 2012; Lawson et al., 2016.
156. European Court of Auditors, 2010; Lawson et al., 2013.
157. Lawson, 2007; Lawson et al., 2016; Rao/Marquette, 2012.
158. Lawson, 2007; Rao/Marquette, 2012; Tidemand, 2009; Williamson/Dom, 2010.
159. Hayman, 2011; Jimat Consult, 2011; Lister, 2006.
160. DRN, 2011; Kremser, 2011; Lawson et al., 2016.
161. Leiderer et al., 2007.
162. AfDB, 2011a; Claussen et al., 2006; de Kemp et al., 2011a; Dijkstra et al., 2012; Frantz, 2004; Koch/Morazan, 2010; Lawson et al., 2003a; Lawson et al., 2014; Lawson et al., 2016; Rao/Marquette, 2012; World Bank, 2009b; World Bank, 20013a.
163. de Kemp et al., 2011a; Frantz, 2004; Lawson et al., 2016; World Bank, 2009b.
164. Lawson, 2007; Lawson et al., 2013; Lawson et al., 2014; Purcell et al., 2006; Rao/Marquette, 2012.
165. Lister, 2006.
166. OECD DAC, 2012.
167. Bartholomew et al., 2006; de Kemp et al., 2011a; DRN, 2011; Frantz, 2004; Harding/Gerster, 2004; Hedger et al., 2010; Koch/Morazan, 2010; Lanser et al., 2006; Lawson et al., 2007; Lawson et al., 2013; Lawson et al., 2014; Lister et al., 2006; ODI/Mokoro, 2009; Smith, 2009; Williamson/Dom, 2010; World Bank, 2013b.
168. Bartholomew et al., 2006; Batley et al., 2006; Bogetic et al., 2015; Claussen et al., 2006; de Kemp et al., 2011a; Dijkstra, 2013; Dijkstra et al., 2012; Dijkstra/Grigsby, 2010; Dom/Gordon, 2011; DRN, 2011; EU, 2014; European Court of Auditors, 2010; Frantz, 2004; Gosparini et al., 2006; Harding/Gerster, 2004; Hedger et al., 2010; Koch/Morazan, 2010; Lanser et al., 2006; Lawson et al., 2003a; Lawson et al., 2007; Lawson et al., 2013; Lawson et al., 2014; Lawson et al., 2016; Lieberon et al., 2004; Lister, 2006; Lister et al., 2006; Loquai/Klavert, 2011; ODI/Mokoro, 2009; Purcell et al., 2006; Smith, 2009; Steffensen, 2010; Tidemand, 2009; Williamson/Dom, 2010; World Bank, 2009a; World Bank, 2009b; World Bank, 2013b; World Bank, 2015a; World Bank, 2015b.
169. Batley et al., 2006; Gosparini, et al., 2006; Kremser, 2011; Lanser et al., 2008; Lawson et al., 2003a; Lister et al., 2006; Loquai/Klavert, 2011; ODI/Mokoro, 2009; Williamson/Dom, 2010.
170. Batley et al., 2006; Bogetic et al., 2015; de Kemp et al., 2011a; Caputo et al., 2013; Dijkstra, 2013; EU, 2014; Faust et al., 2012; Hedger et al., 2010; Koch/Morazan, 2010; Lawson, 2007; Lawson et al., 2003a; Lawson et al., 2013; Lawson et al., 2014; Lawson et al., 2016.
171. Bogetic et al., 2015; de Kemp et al., 2011a; Koch/Morazan, 2010.
172. ADE, 2015; Harding/Gerster, 2004; Lawson, 2007; Lawson et al., 2003a; Lawson et al., 2003b; Lawson et al., 2007; Lister et al., 2006; ODI/Mokoro, 2009; Swedlund, 2014; Williamson/Dom, 2010.
173. ADE, 2015; Lister et al., 2006.
174. Harding/Gerster, 2004; ODI/Mokoro, 2009; Williamson/Dom, 2010.
175. Harding/Gerster, 2004; Lawson et al., 2007; Swedlund, 2014.
176. Tavakoli/Smith, 2013.
177. Booth et al., 2005; Claussen et al., 2006; European Court of Auditors, 2010; Frantz, 2004; Gerster/Chikwekwe, 2007; Harding/Gerster, 2004; Koch/Morazan, 2010; Molt/Wahlers, 2008; Swedlund, 2014.
178. Booth et al., 2005; Gerster/Sawadogo, 2003; Harding/Gerster, 2004; Koch/Morazan, 2010.
179. Bartholomew et al., 2006; Batley et al., 2006; Bogetic et al., 2015; Caputo et al., 2013; de Kemp et al., 2011a; Dom/Gordon, 2011; Gerster/Chikwekwe, 2007; Gerster/Mutakyahwa, 2006; Gerster/Sawadogo, 2003; Gosparini et al., 2006; Harding/Gerster, 2004; Hedger et al., 2010; Lanser et al., 2006; Lawson et al., 2003a; Lawson et al., 2007; Lawson et al., 2013; Lawson et al., 2014; Loquai/Klavert, 2011.
180. ADE, 2015; Booth et al., 2005; Caputo et al., 2013; EU, 2014; European Court of Auditors, 2010; Frantz, 2004; Gerster/Chikwekwe, 2007; Lanser et al., 2006; Molt/Wahlers, 2008; Swedlund, 2014.
181. Bogetic et al., 2015; de Kemp et al., 2011a; Molt/Wahlers, 2008; Swedlund, 2014.
182. Harding/Gerster, 2004.
183. Loquai/Klavert, 2011.

184. Lawson et al., 2007; Lawson et al., 2013; Lawson et al., 2014.
185. Batley et al., 2006; Bogetic et al., 2015; Molt/Wahlers, 2008.
186. de Kemp et al., 2011a; Loquai/Klavert, 2011; Molt/Wahlers, 2008.
187. Lawson et al., 2007; Lawson et al., 2014; Molt/Wahlers, 2008.
188. Bogetic et al., 2015; Claussen et al., 2006; de Kemp et al., 2011a; Dijkstra, 2013; Frantz, 2004; Lister et al., 2006.
189. de Kemp et al., 2011a.
190. Batley et al., 2006; Bogetic et al., 2015; de Kemp et al., 2011a; Dijkstra, 2013; Dijkstra et al., 2012; Fernandes Antunes et al., 2010; Harding/Gerster, 2004; Lanser et al., 2006; Lawson, 2007; Lawson et al., 2003a; Lawson et al., 2007; Lawson et al., 2014; Lawson et al., 2016; Leiderer/Faust, 2012; Lister, 2006; Lister et al., 2006; Molenaers et al., 2010; Purcell et al., 2006; Smith, 2009; Steffensen, 2010; World Bank, 2009a; World Bank, 2009b; World Bank, 2013b; World Bank, 2015b.
191. Claussen et al., 2006; Dijkstra, 2013; Dijkstra et al., 2012; Faust et al., 2012; Gosparini et al., 2006; Hayman, 2011; Lawson, 2007; Molt/Wahlers, 2008; Purcell et al., 2006; World Bank, 2009c.
192. Frantz, 2004; Harding/Gerster, 2004; Lister et al., 2006; Tidemand, 2009; Williamson/Dom, 2010.
193. Harding/Gerster, 2004.
194. Lister et al., 2006.
195. Batley et al., 2006; Caputo et al., 2013; Frantz, 2004; Lawson et al., 2011; Lister et al., 2006; Molt/Wahlers, 2008; Smith, 2009; Steffensen, 2010; World Bank, 2015b.
196. Bogetic et al., 2015; DRN, 2011; Gosparini et al., 2006; Hedger et al., 2010; Purcell et al., 2006; Swedlund, 2014; Tidemand, 2009; World Bank, 2013b.
197. OECD DAC, 2012.
198. Lawson et al., 2003a; OECD DAC, 2012.
199. OECD DAC, 2012.
200. Nilsson, 2004.
201. ADE, 2015; Bernardi, 2015; Bogetic et al., 2015; Dijkstra, 2013; Dom/Gordon, 2011; DRN, 2011; Frantz, 2004; Lanser et al., 2006; Lawson et al., 2011; Lawson et al., 2013; Lawson et al., 2014; Lawson et al., 2016; Leiderer/Faust, 2012; Lister, 2006; Molt/Wahlers, 2008; World Bank, 2008; World Bank, 2009a.
202. Molt/Wahlers, 2008; Purcell et al., 2006.
203. Dijkstra et al., 2010; Lawson, 2016.
204. Claussen et al., 2006; Lanser et al., 2006; Lawson, 2016; Lawson et al., 2007; Lawson et al., 2014; Lawson et al., 2016; Lawson/Kipokola, 2013; Purcell et al., 2006.
205. Lawson/Kipokola, 2013; Leiderer/Faust, 2012; Lister et al., 2006.
206. Booth et al., 2005; de Kemp et al., 2011a; Dijkstra, 2013; Dijkstra et al., 2012; DRN, 2011; Frantz, 2004; Lawson et al., 2007; Lawson et al., 2011; Lawson et al., 2013; Lawson et al., 2014; Lawson et al., 2016; Molt/Wahlers, 2008.
207. Lanser et al., 2006; Lawson, 2016; Lawson/Kipokola, 2013.
208. Lawson, 2016; Leiderer/Faust, 2012.
209. de Kemp et al., 2011a; Dijkstra, 2013; Lawson et al., 2007; Lawson et al., 2013; Lawson et al., 2016; Molt/Wahlers, 2008.
210. Booth et al., 2005; Claussen et al., 2006; Dijkstra et al., 2010.
211. ADE, 2015; Claussen et al., 2006; Purcell et al., 2006.
212. ADE, 2015; Batley et al., 2006; Booth et al., 2005; Caputo et al., 2013; DRN, 2011; Lanser et al., 2006; Lawson et al., 2007; Purcell et al., 2006; Thunnissen/Morillon, 2014; World Bank, 2008; World Bank, 2009a; World Bank, 2009b; World Bank, 2014.
213. ADE, 2015; Bartholomew et al., 2006; Batley et al., 2006; Booth et al., 2005; Claussen et al., 2006; de Kemp et al., 2011a; Dijkstra, 2013; Dijkstra et al., 2012; Dijkstra/Grigsby, 2010; Frantz, 2004; Gosparini et al., 2006; Lanser et al., 2006; Lawson et al., 2003a; Lawson et al., 2005; Lawson et al., 2007; Lawson et al., 2011; Lawson et al., 2013; Lawson et al., 2014; Lawson et al., 2016; Leiderer/Faust, 2012; Lister et al., 2006; Loquai/Klavert, 2011; Molt/Wahlers, 2008; Caputo et al., 2013; Purcell et al., 2006; World Bank, 2009a; World Bank, 2009b; World Bank, 2013b; World Bank, 2015a.
214. Batley et al., 2006; Claussen et al., 2006; Gosparini et al., 2006; Lawson et al., 2003a; Lawson et al., 2007; Lister, 2006.
215. Frantz, 2004; Lawson et al., 2014; Lawson et al., 2016.
216. ADE, 2015; Claussen et al., 2006; Dijkstra/Grigsby, 2010; Gosparini et al., 2006; Lawson et al., 2014; Lawson et al., 2016; Molt/Wahlers, 2008; Purcell et al., 2006.
217. Bartholomew et al., 2006; Caputo et al., 2013; de Kemp et al., 2011a; Dijkstra et al., 2013; Lawson et al., 2011; Lawson et al., 2016; Lister et al., 2006; World Bank, 2015a.
218. Booth et al., 2005; de Kemp et al., 2011a.

219. de Kemp et al., 2011a; Dijkstra/Grigsby, 2010; World Bank, 2009b; World Bank, 2009c.
220. Booth et al., 2005; de Kemp et al., 2011a; World Bank, 2009b; World Bank, 2009c.
221. Claussen et al., 2006; Gosparini et al., 2006.
222. Booth et al., 2005; de Kemp et al., 2011a; Dijkstra et al., 2012; Dijkstra/Grigsby, 2010; Frantz, 2004; Hedger et al., 2010; Lanser et al., 2006; Lawson, 2007; Lawson et al., 2007; Lawson et al., 2011; Lawson et al., 2013; Lawson et al., 2014; Lawson et al., 2016; Molt/Wahlers, 2008; Thunnissen/Morillon 2014; Williamson/Dom, 2010; World Bank, 2009a; World Bank, 2009c; World Bank, 2013b.
223. Alavuotunki, 2015; Bogetic et al., 2015; Caputo et al., 2013; de Kemp et al., 2011a; Dijkstra et al., 2012; Dijkstra/Grigsby, 2010; Lawson et al., 2014; Lanser et al., 2006; Lawson, 2007; Lawson et al., 2011; Lawson et al., 2016; Molt/Wahlers, 2008; Smith, 2009; Thunnissen/Morillon 2014; Williamson/Dom, 2010; World Bank, 2009a; World Bank, 2009c; World Bank, 2013b.
224. Batley et al., 2006; Caputo et al., 2013; DRN, 2011; Frantz, 2004; Hedger et al., 2010; Lawson et al., 2011; Lawson et al., 2013; ODI/Mokoro, 2009; Smith, 2009; Steffensen, 2010; Tidemand, 2009; World Bank, 2008.
225. AfDB, 2011a; Hedger et al., 2010; Lieberson et al., 2004; Purcell et al., 2006; Thunnissen/Morillon, 2014; Williamson/Dom, 2010.
226. ADE, 2015; Booth et al., 2005; de Kemp et al., 2011a; Dijkstra et al., 2012; DRN, 2011; Hedger et al., 2010; Lanser et al., 2006; Lister et al., 2006; Molt/Wahlers, 2008; Thunnissen/Morillon 2014; Williamson/Dom, 2010; World Bank, 2008; World Bank, 2009a; World Bank, 2013a.
227. World Bank, 2009c.
228. Hedger et al., 2010.
229. Williamson/Dom, 2010.
230. Batley et al., 2006; Bogetic et al., 2015; Caputo et al., 2013; Gosparini et al., 2006; Harding/Gerster, 2004; Hedger et al., 2010; Lawson, 2007; Lawson et al., 2011; Lawson et al., 2016; ODI/Mokoro, 2009; Smith, 2009; Thunnissen/Morillon 2014; Tidemand, 2009; Williamson/Dom, 2010.
231. Bogetic et al., 2015; Lawson et al., 2007; Lawson et al., 2013; Lawson et al., 2016; Purcell et al., 2006; Thunnissen/Morillon, 2014; Williamson/Dom, 2010.
232. Lawson 2007; Lawson et al., 2016; Thunnissen/Morillon, 2014.
233. Williamson/Dom, 2010.
234. de Kemp et al., 2011a; Frantz, 2004; Lawson et al., 2014; Molt/Wahlers, 2008; Thunnissen/Morillon, 2014; World Bank, 2013b; World Bank, 2015a; World Bank, 2015b.
235. ADE, 2015; Batley et al., 2006; Caputo et al., 2013; Hedger et al., 2010; Lawson et al., 2007; Lawson et al., 2013; Lawson et al., 2016; ODI/Mokoro, 2009.
236. Claussen et al., 2006; World Bank, 2013b.
237. Hedger et al., 2010.
238. de Kemp et al., 2011a; Lanser et al., 2006; Purcell et al., 2006.
239. Lister, 2006.
240. Williamson/Dom, 2010.
241. DRN, 2011; Hedger et al., 2010; Lawson et al., 2011; World Bank, 2008.
242. Bogetic et al., 2015; de Kemp et al., 2011a; Hedger et al., 2010; World Bank, 2009c.
243. Lister, 2006.
244. Bogetic et al., 2015.
245. Koch et al., 2016.
246. See, for instance: Faust/Koch, 2013; Handley, 2009.
247. EC, 2008b

German Institute for
Development Evaluation (DEval)

Fritz-Schäffer-Straße 26
53113 Bonn, Germany

Phone: +49 228 24 99 29-0
Fax: +49 228 24 99 29-904

E-mail: info@DEval.org
www.DEval.org



DEval

GERMAN
INSTITUTE FOR
DEVELOPMENT
EVALUATION
