WIDER Working Paper No. 2013/105

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

Nathalie Holvoet and Liesbeth Inberg*

October 2013

Abstract

In order to correct for the initial gender blindness of the Paris Declaration and related aid modalities as general and sector budget support, it has been proposed to integrate a gender dimension into budget support entry points. This paper studies the effectiveness of (joint) gender working groups and the integration of sex-disaggregated indicators and targets in performance assessment frameworks in the context of education sector budget support .../

*both authors, University of Antwerp, Institute of Development Policy and Management, corresponding author email: Nathalie.holvoet@ua.ac.be

delivered to a sample of 17 Sub-Saharan African countries over the period 2005-10. Findings of the qualitative comparative analysis demonstrate that engendering these two budget support entry points contributed to high performance on increasing female enrolment.

1   Introduction

Since the turn of the century, changes in aid policies and instruments have taken place with the goal of promoting aid, and ultimately development, effectiveness. The reform agenda which is most clearly laid down in the 2005 Paris Declaration (PD) centres around principles of ownership, harmonization, alignment, managing for results, and accountability. In line with the PD principles, donors have replaced their traditional projects with more programme-oriented aid and budget support. One of the issues that was, at least initially, largely neglected is the gender dimension. Over time, the level of discussion around the gendered implications of the PD has increased (see e.g., Eyben 2007; Gaynor 2007; Holvoet and Inberg, *forthcoming*). The existing research has thus far mainly concentrated on the gender-sensitivity of recipient countries' poverty reduction strategy papers (PRSPs) (see e.g., Bell 2003; Holvoet 2010; Van Staveren 2008). While this focus is understandable from the perspective of country ownership, donors also retain their responsibility with respect to the gender-sensitivity of their aid policies and budgets. From this vantage point, a number of measures have been suggested that donors can introduce to increase the gender-sensitivity of sector and general budget support. These measures include, among others, the inclusion of sex-disaggregated or gender-related indicators and targets in performance assessment frameworks (PAFs) which are important instruments during joint (sector) reviews and disbursement decisions and the set-up of joint gender (sector) working groups that bring together gender experts from various settings (government, donors, civil society) and that lobby and participate in joint (sector) reviews and policy dialogues (see e.g., EC/UN Partnership on Gender Equality for Development and Peace 2008; Holvoet 2010; OECD/DAC 2008a). While in a number of countries these measures have been implemented in the context of (sector and general) budget support, to the best of our knowledge there is little to no systematic research with respect to the use and effectiveness of these measures on the ground. Our research aims, in particular, to fill this gap in the aid and gender literature and will explore whether these measures have contributed to changes in gender-specific outcomes, in what ways and under which circumstances.

We link our research to the ongoing discussion in the gender mainstreaming literature regarding the comparative effectiveness of different types of incentives to trigger the implementation of gender mainstreaming. In line with the gender mainstreaming literature, we categorize the use of indicators and targets in PAFs as a 'harder' type of incentive as compared to the use of the more 'softer' incentive of gender co-ordination networks/working groups. While some (see e.g., Hafner-Burton and Pollack 2009; Pollack and Hafner-Burton 2010) believe that 'harder' incentives which establish binding and enforceable rules work, others (see e.g., Weaver 2008, cited in Pollack and Hafner-Burton 2010) are more sceptical and stand in favour of softer incentives that emphasize socialization and persuasion. If anything, there is thus far relatively little (comparative) empirical research on this topic (see Hafnor-Burton and Pollak 2009; Pollack and Hafner-Burton 2010), which also makes our study interesting from this additional perspective.

Our research focuses on budget support to the education sector, including sector budget support (SBS) and general budget support (GBS) with an education focus, as this is one of the key budget support sectors where over the past decade sex-disaggregated indicators have been introduced in PAFs and in which gender working groups have also been operational. Second, due to their prominent presence in education budget support operations in a wide range of developing countries, we focus on countries where the European Community (EC) was

involved as a donor. Third, in order to increase internal validity, we have opted to increase sample homogeneity by focusing on one specific region, i.e. Sub-Saharan Africa (SSA). Furthermore, it is particularly SSA which is lagging behind with respect to girls' education outcomes such as net enrolment ratio (NER) or completion, while there is also enough variation in outcomes to allow for useful cross-country analysis. Given the fact that aid-related measures are not implemented in a social vacuum but in countries with a specific gender, education, and development context, we also include variables that are proxies for these different contexts. Data collection draws upon existing international databases, a review of country-based aid documentation combined with insights from previous policy advisory research on gender and aid effectiveness for different aid agencies (EC, Belgium, Netherlands, ILO/ITC, UN Women), as well as (limited) input from a questionnaire sent to EC gender focal points. As we are interested in exploring multiple pathways that might lead to the same outcomes, we use qualitative comparative analysis (QCA) (Rihoux and Ragin 2009).

Findings highlight that the inclusion of sex-disaggregated NER indicators in PAFs and/or the presence of a gender working group contributed to high performance in increasing female NER. In countries with a supportive context (free education in combination with a relatively low gender discrimination and/or high aid dependency), the presence of a gender working group has been sufficient; in countries with a less supportive context the inclusion of sex-disaggregated NER indicators has been necessary.

The remainder of the paper is organized as follows. Section 2 presents the main building blocks that guide the empirical research while Section 3 introduces issues related to research methodology. Findings of the analysis are presented in Section 4. Section 5 discusses policy implications while Section 6 concludes.

2 Exploring gender effectiveness of education budget support: towards an analytical framework

No analytical framework for the topic under study was readily available, so we first invested in the development of a framework that would guide the empirical research. Our topic is at the nexus of different streams of literature, including research on gender and changing aid modalities, gender and education, and use of hard and soft incentives in gender mainstreaming and education aid effectiveness. In what follows we give a brief summative account of these different strands of literature and mainly focus on issues that are considered relevant from our own research perspective.

2.1 Changing aid modalities through a gender lens

While the rationale for a gender sensitive PD may easily be built upon equality, effectiveness and efficiency arguments, the gender dimension was largely neglected in the PD, which contains only one passing reference to gender issues in the paragraph on harmonization (OECD/DAC 2005: 7). Thanks to the joined efforts of femocrats (e.g., OECD/DAC Gendernet) and gender actors in civil society, the Accra Agenda for Action (3rd High Level Forum on Aid Effectiveness 2008) and the Busan Partnership Agreement (4th High Level Forum on Aid Effectiveness 2011) pay slightly more attention to gender issues. While there is generally agreement about the opportunities and risks that the five key PD principles in theory hold for gender equality and women's empowerment (see e.g., EC/UN Partnership on Gender

Equality for Development and Peace 2008; Eyben 2007; Gaynor 2006; 2007; Social Development Direct et al. 2008; Van Reisen and Ussar 2005), there is much less evidence with respect to what actually happens on the ground. There is some research on the gender-sensitivity of the PRSPs (see Bell 2003; Holvoet 2010; Van Staveren 2008) and additionally the optional gender module[1] included in the 2011 PD survey gives us some insight into the recipient part of the reform agenda though large-scale joint processes and impact evaluations on budget support have remained largely silent with respect to gender issues (Wood et al. 2008; Wood et al. 2011). Findings of the 2011 optional gender module highlight that since 2005 most progress has been made in addressing gender equality in national development strategies, especially in the areas of social or human development. More specifically, three out of 24 countries reported that gender equality and women's empowerment were to a large extent mainstreamed in their national and sector policies, while another 17 countries highlighted that gender issues were to some extent taken on board. However, as most of the gender priorities are not budgeted for the risk of policy evaporation is high. Countries scored less on the 'gender equality results' and 'mutual accountability for gender equality' indicators. Only half of the countries address gender equality issues during broad-based country dialogues while only two countries reported that significant progress was made on the availability and use of sex-disaggregated data. In another 12 countries, disaggregated data on key monitoring indicators is available but not yet systematically. Finally, the survey findings highlight that where sex-disaggregated indicators have been used, they have led to an increased focus on, and budget allocation for, gender equality and women's empowerment (OECD 2011).

If anything, there is not much specific tracking of the way in which donor agencies are dealing with gender equality and women's empowerment objectives in the context of changing aid modalities. The OECD/DAC data on the gender equality policy marker (G-marker)[2] provide us with interesting information with respect to the gender-sensitivity of OECD/DAC donors' overseas development assistance (OECD/DAC, CRS database),[3] yet the G-marker is not applied to GBS and scoring is done on the basis of project documentation which does not always capture implementation realities on the ground.[4]

While the locus of attention on the recipients' side rightly reflects the shift in responsibilities propagated under the changing aid architecture, more attention for how donors are dealing

---

[1] The gender module includes three gender equality indicators: (i) gender equality and women's empowerment are grounded in a systematic manner in national development strategies (ownership); (ii) data are disaggregated by sex (managing for gender equality results); (iii) mutual accountability for gender equality and women's empowerment (OECD/DAC, 2010: 4).

[2] The G-marker was introduced in 1997 with the aim to facilitate monitoring and co-ordination of DAC member´s activities in support of gender equality. Three different values are used in the marking system: principal policy objectives (G-2) are those which can be identified as being fundamental in the design and impact of the activity and which are an explicit objective of the activity. Significant policy objectives (G-1) are those which, although important, are not one of the principal reasons for undertaking the activity. Not targeted to the policy objective (G-0) means that the activity has been screened against, but was found not be targeted to, the policy objective (OECD/DAC, 2008b: 2).

[3] In 2011, 4.65 per cent and 26.24 per cent of bilateral sector allocable aid that was screened (67.92 per cent of total bilateral sector allocable aid was screened) was allocated to aid with gender equality as principal and significant objective, respectively. In 2002 and 2006 these figures amounted to 2.65 per cent and 23.97 per cent (52 per cent of bilateral sector allocable aid screened) and 2.50 per cent and 23.75 per cent (55.76 per cent of bilateral sector allocable aid screened), respectively (OECD/DAC, CRS database).

[4] Such ex-ante scoring is likely to be an overestimation of the actual situation on the ground given the evidence with respect to the dilution of the gender dimension throughout the intervention cycle.

with gender concerns in the realm of these ongoing changes is just as relevant. As historical records point out, changes in gender/women and development policies have always largely been shaped by changes in general aid policies (see Molyneux and Razavi 2005; Razavi and Miller 1995). From this vantage point, it is particularly interesting to refer to the 2007 OECD/DAC study which highlighted that many donor agencies do not really know how to handle gender issues in sector and general budget support (OECD/DAC 2007); a finding which has been supported by the 2012 African Development Bank synthesis review of 26 thematic and country evaluations of gender mainstreaming in bilateral and multilateral donor agencies (African Development Bank 2012).

In response to the 2007 study, useful suggestions have been made from within the OECD/DAC (2008a) and UN Women (2010). In brief, the key message is that the entry points which donors generally use in budget support operations to influence a country's policies and systems are equally well suited to promote the inclusion of a gender dimension in national policies and systems. Entry points that are commonly used include ex-ante appraisal of policies and systems, mapping of non-state actors, policy dialogue, capacity building, the identification of PAFs, organization of (sector) working groups, monitoring, and evaluative exercises such as joint (sector) reviews. Table 1 gives a brief overview of the budget support entry points and the suggestions made for engendering each of these entry points.

To the best of our knowledge, there is no systematic account of the degree to which these suggestions have been taken on board in the context of budget support operations and little to no systematic evidence with respect to their effectiveness. One noteworthy exception is the 2011 EC evaluation of budget support operations (cases of Mali, Zambia, Tunisia) which highlights that in Mali the disaggregation of health and education indicators in PAFs and SBS matrices and the inclusion of gender issues during health and education sector policy dialogues has led to an increase of girls' enrolment and completion rates in primary education and their progression rates in secondary education, while gender parity remained unchanged. Similarly in Zambia girls' enrolment improved while dropout rates remained problematic at the higher basic level, with gender gaps being particularly high in the poorest and rural areas (Caputo et al. 2011).

Looking at the different policy suggestions made in Table 1, we can easily distinguish among 'harder' incentives such as the inclusion of indicators and targets in PAFs and 'softer' incentives such as the organization of gender (sector) working groups where donors and country stakeholders harmonize amongst each other with respect to joint analytical work, joint priority setting, etc. In doing this, we conceptually link to the ongoing discussion in the gender mainstreaming literature around the use of 'hard' and 'soft' incentives.

Table 1: Engendering budget support entry points

Budget support entry points	Integration of a gender dimension
Ex-ante analysis of a country's policies and systems to diagnose existing strengths and weaknesses	Gender scan of a country's policies and systems, including e.g., gender-aware policy appraisals and institutional analysis of gender-sensitivity of country's institutional apparatus Verification of whether existing national gender policies are integrated into national and sector development policies and plans Verification of whether the existing gender apparatus is involved in policy-making, planning, budgeting, implementation, monitoring and evaluation
Mapping of national non-governmental actors (including CSOs, research institutes, parliament)	Map and assess the capacity of the national gender demand side Assess the extent to which the national gender demand side is involved in national development and sector processes and systems
Capacity development to tackle weaknesses identified during the diagnosis	Integrate gender issues in mainstream capacity building Invest in a country's own gender expertise inside and outside government
Use of PAFs to jointly identify targets that are used as a basis for monitoring and review and disbursement decisions	Integrate jointly identified sex-disaggregated and gender targets in PAFs (preferably selected from country's sector and/or gender policies)
Set up of joint mechanisms for continuous dialogue among donors and country, including e.g., joint (sector) working groups	Set up of joint gender (sector) working groups
Organization of evidence-based policy dialogue	Put gender issues on the agenda of evidence-based policy dialogue
Set up of joint monitoring and review mechanisms, including joint (sector) reviews	Include gender issues during monitoring and evaluation exercises (as dependent and independent variable), including during joint (sector) reviews

Source: based on Holvoet (2010), OECD/DAC (2008a), UN Women (2010).

2.2 Hard and soft incentives in gender mainstreaming

Since its introduction at the 1995 Beijing conference, gender mainstreaming has become a much discussed buzzword in gender research and policy. While most countries in the world have at the least a discourse on gender mainstreaming, research highlights that in reality gender mainstreaming often evaporates throughout the intervention cycle (Holvoet 2010; World Bank 2001). Factors that have been put forward to explain this phenomenon of policy

evaporation include, among others, the lack of political willingness and senior management leadership, the unclear division of responsibilities between gender and non-gender expertise, the lack of budgets and capacities, deficient monitoring and evaluation, and the lack of accountability structures and clear-cut incentives (African Development Bank 2012; Hafner-Burton and Pollack 2009; Moser and Moser 2005; Pollack and Hafner-Burton 2010). The discussion with respect to incentives has lately centred on the use of 'hard' and/versus 'soft' incentives. Hard incentives refer to binding and enforceable measures while soft incentives are less precise, non-binding, and dependent on voluntary compliance (Pollack and Hafner-Burton 2010). Pollack and Hafner-Burton (2010: 292) highlight that 'in reality there is more of a continuum, from soft to hard, including for instance the establishment of coordinating committees and networks of gender mainstreaming officials, the collection of sex-disaggregated statistics, checklists, manuals, handbooks, gender training, gender impact assessment of policies, post-hoc monitoring and evaluation to the enforcement through positive and negative sanctioning of public officials'.

In their comparative case study research on the effectiveness of gender mainstreaming in the EC, Pollack and Hafner-Burton (2010) demonstrate that cross-cutting mandates such as gender and environment are more successfully implemented when hard incentives are being used. Others (Weaver 2008 cited in Pollack and Hafner-Burton 2010) however are more sceptical and point out that hard incentives mainly provoke ritual compliance rather than far-reaching changes in social institutions.

One set of 'harder' incentives that is particularly relevant in the context of changing aid modalities and the more general move towards results-based management and budgeting is the use of indicators and targets, amongst others, in PAFs that are used for monitoring and evaluation (M&E) and disbursement decisions. The use of targets in general and gender targets in particular is not without discussion. While a shift in focus from inputs towards results is generally welcomed and might tackle, or at the least unveil the phenomenon of policy evaporation and function in itself as an incentive, well-known side effects such as gaming[5] and crowding-out[6] have also been documented (see e.g., Kalk et al. 2010). In their analysis of the effect of benchmarking on closing down gender gaps in the European Union, Van der Vleuten and Verloo (2012) point out similar 'reactivity' effects of targets and benchmarks which incentivize the improvement of the quality of documentation on gender inequality rather than stimulating progress in closing down gender inequality itself. According to Lewin (2005) side-effects are also prevalent in the education sector, where the strong focus on primary education and gender equality in primary education has led to less attention for secondary education. Additionally, with regards to vague notions such as gender equality and women's empowerment, targets and indicators are not easily defined and agreed upon. Furthermore, when 'management for results' is misconceived as 'management by results', there is a tendency to focus on short-term deliverables and quick wins which generally invokes a bias towards the education and health sector at the expense of the more deep-rooted gender concerns in the area of time and task allocation (see also Eyben 2010). It has also been noted that results-based management often leads to 'monitoring and evaluation' being narrowed down to pure monitoring and stocktaking (see White 2002) whereas it is the more analytical evaluative exercises in particular that are important from the perspective of unveiling the underlying 'gendered' causes of the observed inequalities in outcomes. Finally,

---

[5] Gaming refers to the falsification of results for maximization of reward (Kalk et al., 2010).

[6] Crowding-out refers to a diminishing or erasing of intrinsic motivation due to external rewards and the focus on indicators that are in the systems thereby neglecting unrewarded indicators (Kalk et al., 2010).

it is also argued that gender aid conditionalities might be counterproductive as they tend to reinforce the idea that gender concerns are foreign ideologies imposed from the outside (Elson and Mc Gee 1995). From this vantage point, the use of 'pro forma or consensual conditionalities' (Killick 1998) referring to the joint identification of targets that are in line with nationally owned policies is particularly important (see also Caputo et al. 2011). With regards to the case at hand, this involves the integration of gender and sex-disaggregated indicators and targets into PAFs that have been used in national sector policies or national gender strategies, amongst others.

In the context of changing aid modalities, there is also an increased use of 'softer' incentives that rather emphasize persuasion and socialization. Specific examples include gender (sector) working groups which facilitate harmonization of gender priorities and expertise of different donor agencies while the presence of national stakeholders from within and outside government also stimulates exchange, discussion and alignment with national gender policies and expertise. Our own field research in Mozambique and Tanzania highlighted that the use of a division of labour approach whereby different members of the gender working group are also members in other (sector) working groups might be particularly useful to mainstream gender throughout the different aid architecture mechanisms. The success of the division of labour approach is of course dependent upon the expertise and track record of the gender working group members, the commitment/openness to gender issues within the other (sector) working groups and the degree of overlap among the frames that are commonly used by gender, sector, public finance management (PFM) experts (Holvoet and Inberg 2008; 2011).

We may also learn from other settings where external incentives in the form of gender targets have been used to feed into our own analytical framework. A case at hand that is relatively well-documented is the setting of gender targets for countries in the context of the EU adhesion process. Sedelmeier (2009) and Montinola (2010), among others, have showcased that for external incentives to be effective, they have to be combined with a number of favourable domestic conditions, including democratic governance and control of corruption while at the same time the credibility of external incentives and use of realistic targets that take into account country realities proves important (see also Lewin 2005). As regards our own study, one straightforward way of localizing indicators and targets is through joint identification within the setting of a joint gender (sector) working group, which hints at the fact that 'hard' and 'soft' incentives might also be mutually reinforcing.

2.3 Effectiveness of education aid

In order to empirically study the comparative effectiveness of different types of incentives, we focus on the education sector. Our sector focus can be justified on multiple grounds. First, within the aid effectiveness literature there is a recent move from a focus on general aid effectiveness to sector aid effectiveness. As causal chains in sector aid effectiveness research are generally less long and complex, there is more room for refinement and more possibilities exist to control for the influence of other intervening variables (see Birchler and Michaelowa 2013; Christensen et al. 2011; Dreher et al. 2008; Michaelowa and Weber 2008). Within this research strand, the education sector has received much attention as donors have devoted substantial shares of aid resources to the education sector in the context of the Millennium Development Goals (MDGs) and the Education for All (EFA) initiative, while the education sector is also at the forefront of the implementation of sector-wide approaches and (sector) budget support.

Our focus on the education sector can also be justified from a gender perspective. As Razavi (1999: 430) puts it, 'education is still considered the gender jewel in the policy crown and the only highly profiled gendered policy description in most of the poverty assessments'. As highlighted above, making 'gender equality' operational in a context of results-based management often leads to a bias towards gender equality in primary and secondary education (see e.g., the MDGs; African Development Bank 2012; Saith 2006). While this 'reductionism' is obviously not without criticism, it simultaneously makes the education sector the most logical—if not, the only possible—choice for research that focuses on the effectiveness of sex-disaggregated indicators and targets. Besides this, there is also a vast gender and education literature (see e.g., Kane 2004; King and Hill 1993; Rugh 2000; Sutherland-Addy 2008; UNESCO 2013; UNGEI 2012) that has identified both demand and supply side factors that influence gender-specific education outcomes and that will be useful inputs in the development of our analytical framework. These factors include among others (household) income, cultural barriers, adult literacy, the gender parity of teaching staff, the cost (direct, indirect, opportunity) of education. While the obstacles to girls' education are clearly diverse, most publications acknowledge the crucial importance of income (at country and household level) and costs on the one hand, and societal institutions (social norms, customs, rights and laws) on the other hand. While indirect costs and opportunity costs also add explanatory power to girls' lower participation in education, it is suggested that direct costs alone are often already an insurmountable obstacle for poor households. This is obvious from research in Malawi, Ghana, and Uganda, for example, which has highlighted the particularly high effects on girls' enrolment rates resulting from an abolishment of user fees (see Deiniger 2003; Kane 2004; Sutherland-Addy 2008; Vandemoortele 2002; World Bank 2009).

While income is a critical factor, it has been demonstrated that societal norms and laws can play at least as important a role, as they shape the context in which households and individuals take decisions with respect to educational investments. More specifically, son preference, early marriage and initiation ceremonies such as female genital cutting, severe restrictions on girls' mobility, limited de-jure and de-facto female ownership rights all contribute to limiting girls' enrolment (Kane 2004; Sutherland-Addy 2008; UNGEI 2012).

As highlighted in Christensen et al. (2011), the academic education aid effectiveness literature itself is still relatively recent without much conclusive evidence or insights into how aid translates into better education outcomes, and at the same time is also largely gender blind. An in-depth review of the literature is beyond the scope of this article, yet we select a number of insights with respect to research methodology and findings that may feed into our own research. While the difference between quality and quantity of education is generally acknowledged and also captured in the EFA goals, education quality has scarcely been addressed in the education aid effectiveness literature, as a result of the lack of large N-databases that adequately capture quality dimensions (Birchler and Michaelowa 2013; Dreher et al. 2008). Some studies have used proxies for quality such as pupil-teacher ratios and literacy rates, even though these indicators are not necessarily linked to improvements in the education system (Birchler and Michaelowa 2013). If anything, most research focuses on enrolment rates, which is considered a credible first step (Christensen et al. 2011) and this holds particularly true for girls' education where entry at primary schooling is conceived as the most important step (UNESCO 2013). Besides enrolment rates, completion rates are often looked at, while distinctions are also made between primary, secondary, and tertiary education. All these levels of disaggregation are also relevant from a gender perspective as findings have demonstrated that the gender gap increases with the level of education, while there is no unequivocal evidence with respect to an increase of gender gaps in completion rates as compared to enrolment rates. While some studies contend that when girls enter

school, they have the same probability as boys to remain in school (UNESCO 2013), others showcase that girls have a higher probability of dropping out for a diversity of reasons, including early marriage, initiation ceremonies, and restrictions on mobility when they enter puberty (Kane 2004). Turning to the aid variables under study, the focus has thus far mainly been on aggregate aid flows, while some studies have also started to explore the potential differential effects of different aid flows (see e.g., Christensen et al. 2011; Michaelowa and Weber 2008).

With regard to findings, one general conclusion, which applies more to the traditional forms of education aid, is that aid for education is more effective than the government's own budget for education. While the latter is mostly spent on salaries, donors put more emphasis on building schools, school management improvement and access for disadvantaged groups (Dreher et al. 2008; Roberts 2003). In line with this is the observation that the availability of financial (aid) resources alone is not enough to influence education outcomes. Certain structural parameters of the education system and the wider institutional setting also need to change in order to reach any of the international education goals. More specifically, education aid seems to be more effective when combined with good governance and reduced corruption (Christensen et al. 2011), and political and institutional governance seem to be particularly important (Michaelowa and Weber 2007). This somehow lends support to those that are in favour of GBS which particularly emphasizes these broader governance reforms, focusing amongst others on issues related to PFM. Others (see e.g., Particip 2010) rather point at the importance of deepening the sector dialogue in order to trigger reforms in sector policies and systems, an area in which SBS has a comparative advantage over GBS.

Building on the above strands of literature we have developed an analytical framework which is visualized in Figure 1. Our main focus of attention is the effectiveness of two different types of gender incentives used by donors to trigger changes in gender education goals while we will also include a number of aid, country and gender context variables which have been distilled on the basis of the above mentioned gender conditionality and education aid effectiveness research. The different variables included in Graph 1 will be made operational in section 3. As we are particularly interested in the possible interplay of the different factors under study, we have employed the QCA for the analysis.

Figure 1: Analytical framework for studying effectiveness of gender-sensitive (non-financial) donor entry points in the context of budget support to the education sector
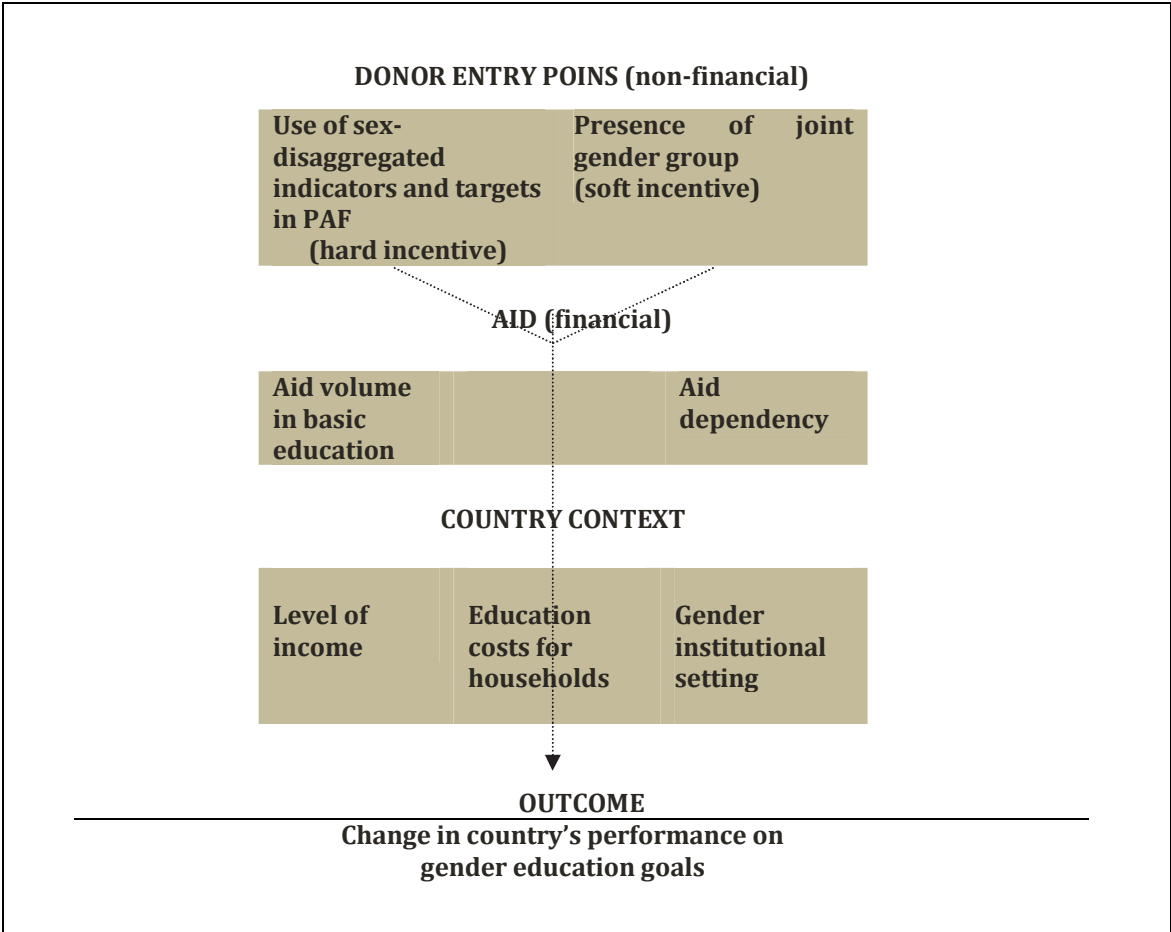
Source: authors' own compilation.

3   Methodology, sample, and data

3.1 Methodology

QCA, developed by Ragin in 1987, was originally presented as a research approach occupying a niche between case-oriented (or 'qualitative') and variable-oriented ('quantitative') methods, reflecting the 'best of both worlds' (Ragin 1987; Rihoux and Marx 2013). In contrast to statistical methods, in which a large number of cases and a relatively small number of variables are used, QCA allows the systematic comparison of a relatively small number of cases and to reduce complex social situations by transforming cases into configurations (Rihoux and Ragin 2009). A configuration refers to a given combination of conditions (e.g., presence of sex-disaggregated indicators in PAFs and presence of gender (sector) working groups) and an outcome (e.g., change in female NER), and can cover one or multiple cases (Rihoux and Lobe 2009). Central in QCA is the identification of (combinations of) conditions that can be interpreted in terms of necessity and sufficiency. A condition is considered necessary if when the outcome is present, the condition always is present. A condition is considered sufficient if when the condition is present, the outcome is always present (Schneider and Wagemann 2010: 3).

While statistical methods like correlation and regression disregard distinct patterns and outliers, QCA broadens the notion of causality by viewing it as context and conjunction specific (Rihoux and Ragin 2009). It recognizes that generally a combination of conditions generates an outcome and that several combinations of conditions can produce the same outcome (equifinality). In addition, the same condition can be sufficient for the presence *and* absence of an outcome if it is combined with other, different conditions. Asymmetrical causality is crucial when using QCA as it implies that the analyses of the occurrence of a phenomenon (e.g., increase in female NER) and non-occurrence of that phenomenon (e.g., no increase in female NER) are to be conducted separately, because the causal mechanisms are not necessarily reversed.

Before proceeding to QCA proper, for which specific software (Tosmana) is used, cases, outcomes and conditions need to be selected and dichotomized, in order to create a dichotomous data table, in which the outcomes and conditions for each case are allocated a '0' or a '1' .[7] Dichotomized outcomes and conditions are obtained through the identification of calibration points. While using hypothesis-driven calibration points is the best option (see Schneider and Wagemann 2010), mechanical cut-off points could be used if this is not possible (Rihoux and Ragin 2009). The process of selection and dichotomization is described in the following two sections.

3.2 Sample

The focus on SSA is based upon the fact that most of the challenges that exist in the area of gender equality in education are concentrated in this region. While progress has been made in SSA, with the gender parity index at primary level increasing from 0.85 in 1999 to 0.93 in 2010, 12 out of the 17 countries with severe gender disadvantage (gender parity index of below 0.90) are situated in SSA (UNESCO 2013: 106). The main reason for fewer girls being in school is that they are less likely to start school. In 2010, the average girls' NER in primary education was 74 per cent for the SSA region, as compared to 78 per cent for low-income countries and 88 per cent as the world average (UNESCO 2013: 355). With regards to the MDG on primary completion, regional disaggregated statistics demonstrate that three fourths of the countries that are the furthest from meeting this education MDG are located in SSA while about half of those are unlikely to meet the gender equality goal (UNGEI 2012: 22). These SSA averages, however, conceal substantial differences among countries in the region, a variation which opens up opportunities for interesting cross-country comparative analysis.

Given that the focus of our study is on the integration of a gender dimension in donor entry points, we first selected those countries within SSA that received education sector or general budget support with a focus on education, in the period under study. This limited our sample to 30 SSA countries.[8] As will be explained in 3.3., our sample has further been reduced by

---

[7] We relied upon crisp-set QCA analysis (csQCA), which only uses dichotomized data. Recently, however, other QCA methods such as multi-value QCA (mvqca) and fuzzy set QCA (fsqca), have been developed in response to criticisms of loss of information due to dichotomization. Fuzzy set QCA is based on a combination of fuzzy set theory and QCA and allows for partial membership in a set. Put differently, instead of allowing only black or white (csQCA), fsQCA can account for different shades of grey (Emmenegger, 2011). As our two main conditions under study (sex-disaggregated indicators in the PAF, presence of joint gender working group) can easily be dichotomized, we opted for csQCA.

[8] This sample includes Benin, Botswana, Burkina Faso, Burundi, Cape Verde, Central African Republic, Chad, Djibouti, Eritrea, Ethiopia, Gabon, Gambia, Ghana, Guinea Bissau, Ivory Coast, Kenya, Lesotho, Madagascar,

restrictions on data availabilities for the different outcome variables and conditions under study.

3.3 Data

The different elements of our analytical framework are operationalized in Table 2. The choice and operationalization of variables has been guided by the literature review in section 2, data availabilities and the limitations with respect to the number of variables set by QCA[9].

Table 2: Operationalization of the analytical framework

	Variables	Operationalization	Calibration point
	OUTCOME VARIABLES	Change in female NER between 2005 and 2010 (ner)[10]	> 10.5% (SSA average) or lower increase, but already > 95% female enrolment in 2005 = 1 ≤ 10.5% and lower than 95% female enrolment in 2005 = 0
		Change in female survival rate to last grade between 2005 and 2009 (sur)	> 3.2% (SSA average) = 1 ≤3.2% = 0
	Donor entry points (non-financial)	Sex-disaggregated indicators and targets in PAF (p)	Present = 1 Absent = 0
C		Presence of a gender working group (g)	Present = 1 Absent = 0
O	Aid (financial)	Total aid to basic education per primary school-age child (average 2005-10) (a)	> US$14.2 = 1 ≤ US$14.2 = 0
N D			
I T		ODA/ gross national income (GNI) (average 2005-10) (o)	> 10% = 1 ≤ 10% = 0
I O	Income	Gross domestic product (GDP)/ capita (average 2005-10) (d)	> US$1400 = 1 ≤ US$1400 = 0
N S	Education costs for households	Presence of free primary education (f)	Present = 1 Absent = 0
	Gender institutional setting	SIGI 2009 (s)	< 0.15 = 1 ≥ 0.15 = 0

Source: authors' own compilation.

From the broad set of possible outcome variables, we have selected two specific MDG education indicators, i.e. 'change in female NER (primary education) over the period 2005–09 and 'change in female survival rate to last grade'[11] over the period 2005–10. While the first indicator focuses on the access girls have to primary education, the second one is related to

---

Malawi, Mali, Mozambique, Namibia, Niger, Rwanda, Senegal, Sierra Leone, Sudan, Tanzania, Uganda, and Zambia.

[9] The proportion of conditions to cases should be less than 0.33 (Marx,2010).

[10] Change in female NER between 2005 and 2010 is calculated as (NER 2010 – NER 2005)/NER 2005*100. Similarly, change in female survival rate to last grade is calculated as (female survival rate 2009 – female survival rate 2005)/female survival rate 2005 * 100.

[11] While the EFA index includes the survival to the fifth grade, this indicator was replaced by the survival to the last grade in the 2008 revized official list of MDG indicators (see http://mdgs.un.org/unsd/mdg/Host.aspx?Content=Indicators/About.htm).

completion. The specific focus on the primary education level is in line with the type of indicators and targets included in the PAFs (see below). Data on these two indicators is collected from the EFA reports and is available for 18 and 20 respectively of the 30 SSA countries which received education sector or general budget support with a focus on education in the period under study. With respect to the calibration of the outcome variables, we have mainly been guided by the SSA average, as this allowed us to have sufficient variation between the cases.[12] This led to a cut-off point being set at 3.2 per cent for the change in female survival rate at last grade between 2005 and 2009 (higher than 3.2 per cent is considered 'high performance on increasing female survival rate to last grade' whereas lower or equal to 3.2 per cent is considered 'absence of high performance on increasing female survival rate to last grade'). As far as change in the NER is concerned, we have used the SSA average of 10.5 per cent while we have also taken into account the take-off position in 2005 in order to correct for the fact that changes in NER are more difficult to realise for countries that already had higher rates initially.

As regards the two main conditions under study, i.e. gender indicators and targets in PAFs and presence of joint gender working groups, we have used dummy variables (0= absence, 1= presence). The assignment of a score draws upon an intensive search of country aid literature which was combined where possible with input from EC gender focal points and other aid officials. Of the countries with information on the NER and 'survival to last grade rate', we were able to track PAFs and the presence of gender working groups for 18 countries. Data in Table 3 showcases that sex-disaggregated enrolment and survival/completion related indicators were included in 8 and 9 of the 18 countries respectively while gender working groups were present in ten of the 18 countries.

---

[12] As SSA generally has a weak performance on education indicators more globally based cut-off points result in insufficient variation between SSA countries to perform QCA.

Table 3: Overview of use of gender indicators in PAFs and presence of joint gender (sector) working groups

Country	Year PAF	Gender indicators in PAF			Presence joint (sector) gender working group (yes/no)
		Sex-disaggregated enrolment related indicators	Sex-disaggregated survival, completion related indicators	Other sex-disaggregated or gender indicators (general or education sector)	
Botswana[13]		None	None	None	No
Burkina Faso	2007-09	Taux Brut de Scolarisation filles; Taux Brut Accès des filles au première classe du primaire (CP1)[14]	None	Adopter la politique nationale genre	Yes
Burundi		None	None	None	No
Cape Verde	2007	None	None	None	No
Ethiopia	2009 (baselines 2007)	NER for grade 1-8 (and gender parity index for grades 1-4)	Primary school completion rate for grade 8 (and Gender Parity Indexes)	None	Yes
Gambia	2008-10	None	None	None	No
Ghana	2008-10	Gender parity index enrolment	None	Conduct DHS survey, including gender disaggregation of indicators, on a timely and regular basis	Yes
Kenya		None	None	None	No
Lesotho	2010	None	None	None	No
Malawi	2007	None	Survival rate in standard 5 by gender	Female literacy rate	Yes
Mali	2007	Taux brut de scolarisation au premier cycle des filles	Taux d'achèvement en 6e année pour l'année N par genre[15]	Taux brut d'accès au première cycle enseignement par genre;	Yes
Mozambique	2007-09	NER of 6 years of age in the 1st grade—girls	Lower primary education (EP2) conclusion rate	PES/OE (Economic and Social Plan/ State Budget) and BdPES (PES implementation report) whereby the actions,	Yes

---

13 For Botswana we looked at the performance indicators included in the 2002-07 and 2008-13 indicative programmes of the EC, as the EC is the only donor providing budget support.

14 'Taux brut de scolarisation filles' translates into 'female gross enrolment rate'; 'Taux brut d'accès à la première classe de primaire filles' equals 'Female gross intake rate at first class of primary school'.

15 'Taux d'achèvement en 6e année pour l'année N par genre' translates into 'Completion rate at grade 6 for boys and girls (year N)'.

				budgets and progress in gender are reflected	
Niger	at least since 2006	Gross enrolment rates for girls	None	None	No
Rwanda	2008	None	Male/female rate in primary school completion	Introduce and mainstream gender budgeting	Yes
Senegal	2008	Taux d'admission à l'école primaire (CI) filles ; taux brut de scolarisation	Taux d'achèvement pour l'élémentaire (filles)	None	Yes
Tanzania	2006	Net primary school enrolment by gender	Transition rate from standard VII to form 1 by gender	None	Yes
Uganda	2009-10	None	Survival to P7 by gender	Literacy and numeracy proficiency (p3 and p6, total and girls), passing Primary Leaving Examinations (PLE) with grades I-III (total, girls); Targeting of spending in MTEF and budget is gender responsive and aligned with NDP	Yes
Zambia	2007-09	Gender parity index grades 5-9	None	Number of girls receiving bursaries in grades 8-9	No

Source: authors' own compilation.

Comparing data availabilities for the two different outcomes and the two main conditions under study resulted in a sample of 15 countries for the analysis of the NER[16] and a slightly different sample of 15 countries for the analysis with respect to the 'survival to last grade'.[17]

With regard to aid (financial) variables we have included the total aid to basic/education per primary school-age child (average over the period 2005-10) and per cent ODA/GNI (average over the period 2005-10). Data with respect to these variables has been collected from the EFA and World Bank databases and was available for the two sub-samples of 15 countries. As calibration points we have used the SSA average of US$14.2 for total aid to basic education per primary school-age child and for ODA/GNI we have opted for a cut-off point of

---

[16] This sample includes Botswana, Burkina Faso, Cape Verde, Ethiopia, Gambia, Ghana, Kenya, Lesotho, Malawi, Mali, Mozambique, Niger, Senegal, Tanzania, and Zambia.

[17] This sample includes Botswana, Burkina Faso, Burundi, Cape Verde, Ethiopia, Lesotho, Malawi, Mali, Mozambique, Niger, Rwanda, Senegal, Tanzania, Uganda, and Zambia.

10 per cent, in line with what has been used by the Overseas Development Institute (2012) to differentiate between 'high aid countries' and 'middle aid countries'.[18]

When it comes to the choice and operationalization of context variables, we were severely restricted by data availabilities. In line with other multi-country aid effectiveness studies (see e.g., Birchler and Michaelowa 2008; Dreher et al. 2008; Michaelowa and Weber 2007) we have opted to use GDP/capita (average over the period 2005-10) as a proxy for a country's general income level. Additionally, we have included the presence of free primary education as a proxy for costs that households face (absence = 0, presence = 1) and the social institutions and gender index (SIGI) to capture gender discrimination in social institutions. SIGI is a composite index that has been constructed under the auspices of the OECD/DAC and that combines five dimensions including 'family code' (parental authority, early marriage, polygamy), 'civil liberties' (freedom of movement, freedom of dress), 'physical integrity' (violence against women, female genital mutilation), 'son preference' (missing women), and 'ownership rights' (women's access to land, bank, property').[19] Its score is between 0 and 1 and the lower the index the lower the level of gender discrimination. Over time there have been some changes in the way SIGI has been constructed (see Branisa et al. 2013) as well as discussion with respect to its validity and reliability.[20] However, the fact that SIGI is the only index that refers to deep-rooted norms, traditions, formal and informal laws that influence (education) outcomes, makes it more apt for our type of research than other indices such as the gender inequality index which include the (education) outcome measures themselves. Data on GDP/capita was collected from the IMF World Economic Outlook Database, SIGI-data was taken from the 2009 OECD/DAC GID database while data on the presence of free primary education draws upon Tomasevski (2006, used in the 2010 EFA Global Monitoring Report), and World Bank and UNICEF (2009). For the GDP/capita and the presence of free primary education, data was available for the two sub-samples of 15 countries, though SIGI was not available for Cape Verde, which decreased the number of countries with sufficient information to 14 for each outcome.

The calibration for the income variable, GDP/capita, was not straightforward, as only two of the countries under study (Botswana and Ghana) have a GDP/capita above the SSA average (US$2078). Therefore, we used the threshold setter in the software program Tosmana (Cronqvist 2006) to identify an 'open space' in the values. On this basis we have opted for US$1400 as cut-off point. For the gender institutional setting indicator, SIGI, we based our calibration on the classification used by the OECD (nd). For SIGI 2009 the OECD uses five categories, countries with high, medium/high, medium, medium/low and low gender discrimination. We have assigned a '0' score to countries with high and medium/high discrimination (SIGI equal or above 0.15) and a '1' score to countries with medium, medium/low, and low gender discrimination (SIGI below 0.15).

---

[18] According to the classification of the Overseas Development Institute ´high aid countries´ are countries with an ODA/GNI above 10 per cent, ´middle aid countries´ are countries with an ODA/GNI between 2 per cent and 10 per cent, ´low aid countries´ are countries with an ODA/GNO between 1 per cent and 2 per cent and ´very low aid countries´ are countries with an ODA/GNI below 1 per cent (Overseas Development Institute 2012).

[19] See Branisa et al. (2009) and Bransina et al. (2013) for a detailed overview and discussion of the construction of SIGI 2009.

[20] Hawken and Munck (2013) provide a useful comparative analysis of the quality of five composite gender indices. While SIGI scores relatively high on conceptual dimensions through the inclusion of indicators that are not readily available elsewhere, it scores relatively low with respect to indicator scales, indicator value assignment, and aggregation procedures.

On the basis of the information available for the two samples of 14 countries, we produced a table with raw data (see Appendix 1) which we dichotomized (1 or 0) on the basis of the identified calibration points (see Appendix 2).

The different conditions have subsequently been employed in the csQCA analysis to explore how the interplay between the non-financial donor entry points, (financial) aid variables, and country context variables lead to high performance on change in primary enrolment and completion rates. The equifinality principle central to the QCA approach allows us to identify various paths of multiple conjunctural conditions for high performance.

4   Findings and analysis

4.1 Necessary conditions for high performance on increasing female primary enrolment and survival to last grade

Before studying the paths to high performance on increasing female NER in primary education, we first perform a 'necessity analysis'. In doing this, we calculate for each individual condition in the dichotomous data table 'the necessity' (see Table 4). A perfectly consistent 'necessity' (value 1) implies that all cases that are member of outcome y are also member of condition x (but cases can be member of condition x, but not of outcome y) (see Schneider and Wagemann 2012). Taking into account a threshold of 0.90 for necessity (see Schneider and Wagemann 2010), findings in Table 4 highlight that no single condition is necessary. However, the presence of sex-disaggregated indicators in the PAF, the presence of a gender working group and total aid to basic education per primary school-age child[21] come close to the threshold of 0.90. More specifically, our findings show that out of nine countries that show a high performance on increasing female NER eight countries included sex-disaggregated indicators in their PAFs.

Table 4: Overview of 'necessity' scores for the different conditions

Condition	NER	Survival
Sex-disaggregated indicators and targets in PAF	0.89	0.63
Gender working group	0.89	0.63
Total aid to basic education per primary school-age child	0.89	0.88
ODA/GNI	0.78	0.63
GDP/capita	0.22	0.33
Presence of free primary education	0.67	0.63
SIGI	0.44	0.50

Source: authors' own compilation.

Comparing the necessity scores between the NER and the survival rate demonstrates that most necessity scores are higher for improvements in female NER, except for GDP/capita and SIGI. The fact that SIGI is more important for high performance on improvements in female completion than for high performance on improvements in the NER does not really come as a

_____

[21] The number of countries that are members of condition aid/primary school child is however rather large (12 out of 14 countries).

surprise as it is more likely that early marriage practices, for instance, have a higher effect on female completion than on female enrolment (see also 2.3).

In order to limit the number of conditions to be included in the QCA analysis we decided to drop the total aid to basic education per primary school-age child, as this variable demonstrates insufficient variation (only two out of 14 countries in the two samples[22] receive less than the SSA average of US$14.2), and additionally the GDP/capita because of lower necessity rates for high performance on increasing female NER and insufficient variation for the 14 countries as regards high performance on increasing female survival to the last grade (only three countries have a higher GDP/capita than the threshold of US$1400[23]). Additionally, the ODA/GNI variable demonstrates insufficient variation for the sample of countries included for the survival outcome.[24] This elimination process finally generates a set of five conditions for the NER outcome and four conditions for the survival outcome.

While we have been able to identify paths to high performance on increase of female NER in primary education (see 4.2), the QCA analysis for the female survival to the last grade did not produce an acceptable result. On the basis of the information on the four conditions (presence of sex-disaggregated indicators in the PAF, presence of a gender working group, ODA/GNI, presence of free primary education and SIGI) in the dichotomous data table (see Appendix 2) we produced a 'truth table' in which all possible combinations of conditions are presented. With the four conditions the possible combinations were 2^4. Of those 16 possibilities, eight were observed, three of which showed the outcome, two did not show the outcome and three showed contradictory outcomes.[25] A contradictory outcome means that countries with the same combination of conditions do not share the same outcome. Schneider and Wagemann (2012) mention three possible strategies to address contradictory cases. The first strategy is to add other conditions. While we tried various combinations with other variables that are often highlighted in literature to have an influence on completion,[26] this did not solve the contradictions, it rather created new ones. A second strategy is to re-specify the definition of the population of interest. In our introduction we refer to three important elements of focus in our research: budget support with a focus on education, the involvement of the EC as one of the donors, and SSA countries. While we have considered looking at budget support operations where the EC was not involved, this either limited our sample of countries too much or created more data collection problems. Changing the first or last element of focus was not an option as this would divert the topic of study too far away from our initial idea. Furthermore, the last possible strategy, i.e. re-specifying the definition, conceptualization, and/or measurement of the outcome or conditions, did not dissolve the contradictions.

---

[22] Ethiopia and Kenya in the NER group and Ethiopia and Uganda in the survival group.

[23] Lowering the threshold of US$1400 is too artificial as the GDP/capita of the countries under study that have a GDP/capita lower than US$1400 do not differ much (a threshold of US$1200 GDP/capita would for instance separate Tanzania with a GDP/capita of US$1297 and Uganda with a GDP/capita of US$1170).

[24] Three countries with a relative high ODA/GNI (Burundi, Rwanda, and Uganda) are in the survival group and not in the NER group, while two countries with a relative low ODA/GNI (Ghana and Kenya) are in the NER group and not in the survival group.

[25] The first contradictory outcome concerns Burundi (outcome 0) and Lesotho (outcome 1), the second concerns Ethiopia (outcome 0), Mozambique (outcome 0), Rwanda (outcome 1) and Uganda (outcome 1), the third one concerns Malawi (outcome 1), Senegal (outcome 1,) and Tanzania (outcome 0).

[26] Factors that were added to the analysis include governance effectiveness, percentage of population under 15, under five mortality rate, pupil teacher rate and adult literacy rate, amongst others.

Therefore, we had to conclude that the differences in outcomes at the level of survival are probably created by additional other conditions that we were not able to identify.

4.2 Paths to high performance on increasing the female net enrolment rate in primary education

For the analyses of the performance on increasing female NER we used Tosmana (Cronqvist 2006) and introduced five conditions: presence of sex-disaggregated indicators in the PAF, presence of a gender working group, ODA/GNI, presence of free primary education and SIGI. As these five conditions are still too many to conduct meaningful QCA analyses for the 14 cases,[27] we decided to perform the QCA analyses for two sets of conditions, one with PAF, gender working group, free education and ODA/GNI and one with PAF, gender working group, free education, and SIGI.[28]

For the first set of conditions, we created a 'truth table' with 16 (2^4) possible combinations, of which nine are observed. Of these nine observed combinations five show the outcome and four do not show the outcome (see Table 5 and Venn diagram in Appendix 3). The remaining seven possible combinations are logical remainders, which means that for these combinations no cases are found.

In the next step the software uses QCA Boolean algebra to minimize the combinations for the occurrence of the outcome. Boolean algebra employs '+' as OR, and '*' as AND. Capital letters indicate the presence of a condition, while lower case letters indicate the absence of a condition. Four conditions were included in the first analysis: P (PAF), G (gender working group), O (ODA/GNI) and F (free education). The QCA output is based on a process of logical minimization[29] (Grofman and Schneider 2009; Ragin 1987), the '->' symbol indicates that the causal configurations to the left are sufficient for the outcome to the right.

POf + PGF + GOF -> NER[30]

---

[27] According to Marx (2010) the proportion of conditions to cases should be less than 0.33 and with five conditions and 14 cases this would be 0.36. In addition, the benchmark table developed by Marx and Dusa (2011) put forward a minimum of 18 cases for 5 conditions as a threshold for generating non-random meaningful conjunctural paths. With four conditions and 14 cases the proportion to cases is 0.29 and the 14 cases are above the minimum number of 13 cases for four conditions (Marx and Dusa 2011).

[28] The analysis with the four conditions led to a contradictory free truth table with the following solution: PGOS+ POSf + PGsF + GOsF -> NER.

[29] The minimization algorithm basically consists of reducing logically redundant factors, i.e. factors that do not contribute to explaining the outcome. In our case for example, the truth table shows that PGOF is sufficient for high performance on increasing female NER (line 1) and that PGOf is also sufficient for high performance (line 2). This hints at the fact that the state of condition f (F or f) is not important; that it adds little explanatory power and may be discarded, which leads to the identification of PGF as a possible path for high performance on increasing female NER.

[30] Logical remainders could be included in the analyses as hypothetical cases to further minimize the outcome. However, as our study is rather exploratory we could not build upon in-depth theoretical knowledge and therefore excluded these logical remainders. With logical remainders the solutions would be: G + Pf.

Table 5: Truth table for outcome NER with PAF, gender working group, ODA/GNI and free education as conditions

Conditions				Outcome	Country
PAF	gender working group	ODA/GNI	free education	NER	
1	1	1	1	1	Ethiopia, Mozambique, Tanzania
1	1	1	0	1	Burkina Faso, Mali
1	1	0	1	1	Ghana, Senegal
0	1	1	1	1	Malawi
1	0	1	0	1	Niger
1	0	1	1	0	Zambia
0	0	1	1	0	Gambia
0	0	0	1	0	Kenya, Lesotho
0	0	0	0	0	Botswana
1	1	0	0	?	Logical remainder
1	0	0	1	?	Logical remainder
1	0	0	0	?	Logical remainder
0	1	1	0	?	Logical remainder
0	1	0	1	?	Logical remainder
0	1	0	0	?	Logical remainder
0	0	1	0	?	Logical remainder

Source: authors' own compilation.

The above formula highlights that three possible combinations lead to high performance on increasing the female NER. The relative importance of these three paths is shown in Table 7. The first combination is the most prominent path (unique coverage of 0.33) and reads as follows: the presence of sex-disaggregated NER indicators in the **P**AF in an aid dependent country (high percentage of **O**DA/GDP) where there is no **f**ree education leads to high performance on increasing the female NER. The three countries in this group are Burkina Faso, Mali and Niger. The second combination that leads to a positive outcome includes countries with **F**ree education where a **P**AF with sex-disaggregated NER indicators and a **G**ender working group are present. Ethiopia, Mozambique, Tanzania, Ghana, and Senegal are in this group. The third combination includes aid dependent countries (high percentage of **O**DA/GDP) with **F**ree education where a **G**ender working group is present. Countries in this group are Ethiopia, Mozambique, Tanzania,[31] and Malawi. This solution suggests that in aid dependent countries hard incentives (PAF) are necessary in the case that there is no free education. If education is free hard incentives are less relevant, but soft incentives are still necessary. If a PAF and a gender working group is present in a country with free education the level of aid dependency is less relevant.

As mentioned in Section 3, due to asymmetrical causality, the analyses of the absence of the outcome should be conducted separately. The solution for the absence of the outcome is:[32]

pgo + gOF -> ner

---

[31] Cases can follow multiple paths towards the outcome (see Schneider and Wagemann 2012), which explains why Ethiopia, Mozambique, and Tanzania are both in the second and third group.

[32] The unique coverage is 0.6 for the first path and 0.4 for the second path.

The first path is followed by Botswana, Kenya, and Lesotho and means that the absence of sex-disaggregated NER indicators in the **p**af and the absence of a **g**ender working group in a country that is not dependent on aid (relatively low percentage of **o**da/gdp) leads to an absence of the outcome.[33] The second path includes two countries (Gambia and Zambia[34]) that are aid dependent (high percentage of **O**DA/GDP) and have **F**ree education but do not have a **g**ender working group.

The four conditions included in the second analysis are: P (PAF), G (gender working group), S (SIGI) and F (free education). For this combination also, nine out of 16 (2^4) possible combinations are observed, of which five show the outcome, four do not show the outcome and seven are logical remainders (see Table 6 and Venn diagram in Appendix 4).

Table 6: Truth table for outcome NER with PAF, gender working group, SIGI and free education as conditions

Conditions				Outcome	Country
PAF	gender working group	SIGI	free education	NER	
1	1	1	1	1	Ghana, Senegal, Tanzania
0	1	1	1	1	Malawi
1	1	0	1	1	Ethiopia, Mozambique
1	1	0	0	1	Burkina Faso, Mali
1	0	0	0	1	Niger
1	0	0	1	0	Zambia
0	0	1	1	0	Kenya, Lesotho
0	0	1	0	0	Botswana
0	0	0	1	0	Gambia
1	1	1	0	?	Logical remainder
1	0	1	1	?	Logical remainder
1	0	1	0	?	Logical remainder
0	1	1	0	?	Logical remainder
0	1	0	1	?	Logical remainder
0	1	0	0	?	Logical remainder
0	0	0	0	?	Logical remainder

Source: authors' own compilation.

The solution, with no use of logical remainders[35] is:

$$PGs + Psf + GSF \rightarrow NER$$

The first path reads as follows: the presence of sex-disaggregated NER indicators in the **P**AF in combination with the presence of a **G**ender working group in a country where gender discrimination is high (a relatively high **s**igi) leads to a positive outcome. Countries that have

---

[33] In Lesotho female NER decreased with 15.7 per cent between 2005 and 2010, in Botswana and Kenya female NER increased slightly (4.8 per cent respectively 5.0 per cent), but below the SSA average.

[34] In Gambia female NER decreased with 13.0 per cent, in Zambia female NER increased (3.4 per cent), but below SSA average.

[35] With logical remainders: G + Pf or G + Sf.

followed this path are Burkina Faso, Mali, Ethiopia, and Mozambique. The second path is followed by three countries (Burkina Faso, Mali, and Niger) and reads as follows: the use of sex-disaggregated NER indicators in the **P**AF in a country where gender discrimination is high and where there is no **f**ree education leads to a positive outcome. The last path is the most prominent (unique coverage of 0.44, see Table 7) and points out that the presence of a **G**ender working group in a country with relative low gender discrimination (relatively low **S**IGI) and with **F**ree education leads to a positive outcome. Ghana, Senegal, Tanzania, and Malawi are countries that follow this path. From this solution we can conclude that in countries with a high gender discrimination, progress in female NER is made by using hard incentives, either in combination with soft incentives or not. The countries that were able to show results with the use of hard incentives only (Burkina Faso, Mali, Niger) are, however, exactly those three countries that follow the most prominent path in the first solution (POf), which hints at the fact that the level of aid dependency was an influence as well. In countries where gender discrimination was already relatively low, soft incentives were sufficient in combination with free education to trigger high performance on increasing the female NER.

The solution for the absence of the outcome is:

$$pgS + gsF \rightarrow ner$$

The first solution, which is followed by Botswana, Kenya, and Lesotho, points out that the absence of sex-disaggregated NER indicators in the **p**af in combination with the absence of a **g**ender working group in a country with relative low gender discrimination (relatively low **S**IGI) leads to a decrease or a lower than SSA average increase in female NER. The second solution means that the absence of a **g**ender working group in a country where gender discrimination is high (relatively high **s**igi) with **F**ree education leads to the absence of the outcome. Gambia and Zambia follow this path. These two paths can easily be combined with the two paths that do not show the outcome in the first set of solutions (pgo + gOF), as these paths are followed by exactly the same countries. In sum, if hard and soft incentives are not used in countries with a relatively low level of gender discrimination that are not dependent on aid, little or no progress is made in female NER. Additionally, if soft incentives are not used in aid dependent countries with a relative high level of gender discrimination little or no progress is made.

Table 7 provides an overview of the relative importance of the different paths for both sets of conditions. While the raw coverage demonstrates how much of the outcome is covered by each path, the unique coverage demonstrates how much of the outcome is covered only by a specific path (Schneider and Wagemann 2012).

Table 7: Raw and unique coverage for paths to increase in female NER

Countries	Donor entry points (harder and softer incentives)	Aid (financial)	Education costs	Gender Institutional Setting	Raw coverage	Unique coverage
Burkina Faso, Mali, and Niger	PAF (P)	ODA/GNI (O)	free (f)		0.33	0.33
Ethiopia, Ghana Mozambique, Senegal, and Tanzania	PAF (P), Gender working group (G)		Free (F)		0.56	0.22
Ethiopia, Malawi , Mozambique, and Tanzania	Gender working group (G)	ODA/GNI (O)	Free (F)		0.44	0.11
Burkina Faso, Ethiopia, Mali, and Mozambique	PAF (P), Gender working group (G)			sigi (s)	0.44	0.22
Burkina Faso, Mali, and Niger	PAF (P)		free (f)	sigi (s)	0.33	0.11
Ghana, Malawi Senegal, and Tanzania	Gender working group (G)		Free (F)	Sigi (S)	0.44	0.44

Source: authors' own compilation.

While the combination of presence of sex-disaggregated NER indicators in the PAF and a gender working group in combination with free education has the highest raw coverage (0.56), the path with the highest unique coverage (0.44) is followed by Ghana, Senegal, Tanzania, and Malawi, countries that have relatively low gender discrimination, free education and a gender working group.

Interestingly, the presence of sex-disaggregated indicators in the PAF and/or gender working group are included in all solutions. In fact the four countries that do not have sex-disaggregated NER indicators in the PAF and no gender working group (Botswana, Gambia, Kenya, and Lesotho, see Table 5.) do not show a high performance on increasing the female NER.[36]

With regard to the use of incentives one can conclude from the table that the use of soft incentives is only sufficient for an increase in female NER if it is combined with free

_____

[36] The fifth country that does not show an increase in female NER (Zambia) has sex-disaggregated indicators in the PAF, but no gender working group. As this country is highly aid dependent and has free education one could have expected a positive outcome. High gender discrimination in this country (SIGI of 0.22) could have contributed to the lack of performance.

education in countries that are either aid dependent or have relatively low gender discrimination. If education is not free, the use of harder incentives is sufficient in aid dependent countries.

5   Discussion and policy implications

The findings of our research confirm the importance of the use of incentives to promote gender mainstreaming. In the countries under study the inclusion of sex-disaggregated NER indicators in the PAF and/or the presence of a joint gender working group contributed to high performance on increasing female NER. The use of incentives proves to be especially effective in highly aid dependent countries, which does not entirely come as a surprise given the relatively higher level of benefit for recipient countries in these settings. While the inclusion of a gender dimension in donor entry points is certainly not without discussion, it is also not necessarily counter to the PD ownership principle, as most of the countries do have country-owned gender equality and empowerment policy objectives as well as an institutional apparatus and actors with a specific mandate towards those objectives. These national or sector gender policies are, however, more often than not neglected in national poverty reduction strategies, development plans and sector policies while national gender expertise, be it at ministerial or sector level, also often tends to be hardly involved in national development policy-making, budgeting, implementation, and M&E. In such settings, the inclusion of a gender dimension in donor entry points is an effective way to give more weight to nationally-owned gender policies and to increase the room for manoeuvre of the existing gender mainstreaming apparatus (see also Holvoet and Inberg 2008; 2011).

In the two countries with a positive outcome that are not highly dependent on aid (Ghana and Senegal) the local context has been supportive towards the effectiveness of targets and gender working groups. Gender discrimination in both countries is relatively low and education is free. It seems that in a supportive context, in countries with free education and relatively low gender discrimination and/or high aid dependency, the use of soft incentives is sufficient. In a less supportive context the use of hard incentives is necessary, but also not sufficient for a positive outcome. Burkina Faso, Mali and Niger for example do not have free education, and gender discrimination in these three countries is also high, but the fact that they are highly aid dependent contributed to an effective use of the donor entry points.

While in only two countries, Mozambique and Ethiopia, the combination of the presence of sex-disaggregated NER indicators in the PAF and the presence of a gender working group has been necessary for a positive outcome, five other countries[37] with a positive outcome have both sex-disaggregated NER indicators in the PAF and a gender working group. This combined effect of hard and soft incentives does not really come as a surprise. The mandate of joint gender working groups often includes the mainstreaming of gender in the formulation, implementation, and M&E of poverty PRSPs[38] and PAFs increasingly rely upon targets and indicators included in PRSPs. Gender working groups could thus (indirectly) stimulate the inclusion of gender and sex-disaggregated indicators in the PAF. Importantly, the discussion among a broad range of stakeholders in joint gender working groups also helps to identify localized indicators and realistic gender and sex-disaggregated targets, which in turn also increases their effectiveness. Reversely, our own field research in Mozambique (see Holvoet

---

[37] Burkina Faso, Ghana, Mali, Senegal, Tanzania.

[38] See e.g., http://www.ptfsenegal.org/groupes/genre.htm for the mandate of the Senegalese gender working group.

and Inberg 2008) demonstrated that once sex-disaggregated and gender indicators are included in PAFs, gender working groups have more leverage to put gender issues on the agenda of (sector) policy dialogues and joint monitoring exercises such as joint (sector) reviews (see Holvoet and Inberg 2008; 2011).

As gender working groups are absent in the five countries in our sample that do not show a positive outcome, an important first entry point for donors to promote the inclusion of a gender dimension in national policies and systems in these countries could be the setting up of a gender (sector) working group. In fact, providing neutral spaces for discussion among actors from various settings (inside and outside government) and with different comparative advantages (see Guijt 2008 in Eyben 2010) also neatly matches the donor's function of brokerage between state and citizens which is particularly critical in countries where relations between state and society are weak or do not exist (see Booth 2011; Unsworth 2009). As gender experts within bilateral and multilateral donor agencies might be particularly well-placed to fulfil such a brokering role, investing in gender expertise and adapting their mandate and location[39] might be preferable to the current tendency to downsize gender expertise in field offices (Holvoet and Inberg 2012).

Gender working groups may also steer analysis of data collected with respect to sex-disaggregated indicators and targets as well as stimulate effective feedback and use of findings. While many countries under study include sex-disaggregated education indicators in their PAFs, little analysis is done into the underlying causes of the observed inequality which leaves the significance of the underlying gender norms veiled (see also Holvoet 2010). This predominance of monitoring over evaluative analysis is to a large extent related to the narrowly conceived manner in which results-based management has thus far been implemented within donor agencies (Liverani and Lundgren 2007). In fact, this general underinvestment in evaluative exercises is particularly relevant from a gender perspective as it tends to leave the effect of gender blindness on development outcomes unexplored. Unveiling and studying gender-differentiated impacts on the ground may, in addition, be particularly important in efforts on the part of non-state actors to hold the government accountable for promises made earlier.

One useful type of such exercises are gender-responsive budgeting initiatives that aim at analysing policies and related budgetary allocations through a gender lens. The potential of gender responsive budgeting has already been extensively documented (see Budlender 2007; www.genderbudgets.org), yet it remains underutilized (UNGEI 2012). When education budgets and service delivery are scrutinized, gender issues generally receive little attention, notwithstanding some useful exceptions (see Demery 2002). Particularly useful gender budget analysis tools in this respect are tools that track the extent to which boys and girls benefit from education budgets at primary, secondary, and tertiary level (sex-disaggregated benefit incidence analysis), that explore the degree of satisfaction of girls and boys as regards the access and quality of schooling (sex-disaggregated beneficiary assessment) and that study the effect of changes in user fees on boys' and girls' enrolment (sex-disaggregated revenue incidence analysis) (see e.g., Vandemoortele 2002). Given the fact there is a public goods element in such exercises which leads to underinvestment (Center for Global Development 2006), it might be particularly worthwhile for donors to jointly invest in such evaluative

---

[39] In response to changes in the aid setting the substance of the work of gender experts has changed, while most donor agencies have not yet amended their organizational apparatus in order to take better advantage of opportunities embedded within aligned aid for promoting the incorporation of a gender dimension into national policies and systems (see Holvoet and Inberg 2011).

exercises. The evidence generated is not only valuable from an accountability perspective, but also particularly useful to feed into policy dialogues at sector level with respect to strategies that increase gender equality of education outcomes. Such rigorous evidence might also give more leverage to gender working groups in sector dialogues and can be a useful basis to formulate relevant and realistic gender and sex-disaggregated indicators and targets. Additional leverage might also be gained from the fact that gender responsive budgeting offers a bridging frame between gender experts and the ministries/departments of planning and finance. As is evidenced in our own field research in Mozambique such bridging frames can also stimulate networking and co-operation among gender focal points and staff from planning and budget departments within sector ministries (Holvoet and Inberg 2008).

Interestingly, reference to gender responsive budgeting is included in the PAFs of Mozambique, Uganda, and Rwanda (see Table 3), which could be considered a form of 'process conditionality'. Over time there has been an intense discussion with respect to the use of results versus process indicators in PAFs (see e.g., Adam and Gunning 2002; Adam et al. 2004). While some argue that process conditionality is counter to country ownership as it determines the route that needs to be followed towards performance (see Adam and Gunning 2002; Adam et al. 2004), others contend that process indicators are more apt to stimulate systemic changes and tie in more closely with the initial PRSP philosophy (see e.g., Booth 2003), while it also allows more close monitoring of the implementation of the reform process.

Finally, the current move within donor agencies towards portfolio approaches in which GBS, SBS, technical assistance, and targeted 'pilot' projects are used in a coherent fashion (Molenaers and Renard 2008), is also particularly relevant from a gender perspective as gender mainstreaming needs a more all-encompassing approach. It allows the elaboration of a multifaceted strategy whereby donor entry points used in the context of budget support can be combined and linked with capacity development projects of the gender machinery as well as support to domestic gender accountability actors.

6   Conclusions and issues for further research

In this research we analyse the effects of the inclusion of sex-disaggregated indicators and targets in PAFs and the set-up of joint gender (sector) working groups, measures that have been suggested to increase the gender-sensitivity of sector and general budget support. We explore whether these measures have contributed to changes in gender-specific outcomes, in what ways and under which circumstances. In doing this, our research also connects with the ongoing discussion on the comparative effectiveness of 'hard' (inclusion of sex-disaggregated indicators and targets) versus 'soft' (set up of gender working group) incentives for triggering gender mainstreaming.

We focus on the education sector and on SSA countries that received budget support from the EC. Drawing upon a review of various relevant streams of literature in the area of gender mainstreaming and gender and education aid effectiveness, our analytical framework selects a 'harder' (indicators in PAF) and a 'softer' (gender working group) type of incentive, two aid dimensions (aid volume in basic education and aid dependency) and three dimensions that represent the country context (income, education costs for households, and gender institutional setting). As we are particularly interested in the possible interplay of the different factors under study, we have employed QCA for the analysis. In this analysis we focus on two specific education outcomes, increase in female NER and increase in female survival rate to

last grade. While the QCA analysis of the first outcome provides a contradictory free solution, we have been unable to solve contradictions in the analysis of the second outcome.

Findings of the QCA analysis with respect to high performance on increasing female NER highlight that the inclusion of sex-disaggregated NER indicators and/or the presence of a joint gender working group contributed to an increase in female NER. In all solutions the presence of sex-disaggregated NER indicators in the PAF and/or the presence of a GWG are included. In countries with a supportive context (free education in combination with a relatively low gender discrimination and/or highly aid dependency), the presence of a gender working group (soft incentive) has been sufficient, in countries with a less supportive context the inclusion of sex-disaggregated NER indicators (hard incentive) has been necessary. While we presume a positive relation between the presence of a gender working group and the inclusion of sex-disaggregated indicators in the PAF (in seven out of nine countries with a positive outcome both sex-disaggregated indicators in the PAF and a gender working group are present), further research is needed to explore the mutually reinforcing influence of gender working groups and the inclusion of sex-disaggregated indicators in the PAF.

The findings of our study tie in closely with the outcomes of a 2012 OECD/DAC study demonstrating that international targets for gender equality in education have been effective in focusing donor efforts on reducing enrolment gaps between girls and boys. According to this study similar global commitments and targets with respect to gender equality in the economic and productive sectors could help to intensify donor efforts in these critical areas where gender-based inequalities often tend to be even more pronounced than in the education sector (see OECD/DAC 2012). In line with this, one could argue in favour of broadening the inclusion of gender and/or sex-disaggregated indicators and targets in PAFs beyond the education (and health) sector to economic and productive sectors. However, as it is highly likely that other aid and context dimensions will be of influence in these sectors, more fine-tuned research that specifically focuses on these sectors is needed.

References

3rd High Level Forum on Aid Effectiveness (2008). 'Accra Agenda for Action'. Accra: 3rd High Level Forum on Aid Effectiveness.

4th High Level Forum on Aid Effectiveness (2011). 'Busan Partnership for Effective Development Co-operation'. Busan: 4th High Level Forum on Aid Effectiveness.

Adam, C., and J. W. Gunning (2002). 'Redesigning the Aid Contract: Donors' Use of Performance Indicators in Uganda'. *World Development*, 30: 2045-56.

Adam, C., G. Chambas, P. Guillaumont, S. Guillaumont Jeanneney, and J. W. Gunning (2004). 'Performance-based Conditionality: a European Perspective'. *World Development*, 32: 1059-70.

African Development Bank (2012). 'Mainstreaming Gender Equality: A Road to Results or a Road to Nowhere?'. Tunis: African Development Bank, Operations Evaluation Department.

Bell, E. (2003). 'Gender and PRSPs: With Experiences from Tanzania, Bolivia, Vietnam and Mozambique'. Bridge Development—Gender Report No. 67. Brighton: Bridge.

Birchler, K., and K. Michaelowa (2013). 'Making Aid Work for Education in Developing Countries. An Analysis of Aid Effectiveness for Primary Education Coverage and Quality'. Working Paper 2013/021. Helsinki: UNU-WIDER.

Booth, D. (2003). 'Introduction and Overview'. *Development Policy Review*, 21 (2): 131-59.

—— (2011). 'Aid Effectiveness: Bringing Country Ownership (and Politics) Back in'. Working Paper 336. London: Overseas Development Institute.

Branisa, B., S. Klasen, and M. Ziegler (2009). 'Background paper: The Construction of the Social Institutions and Gender Index (SIGI)'. Goettingen: University of Goettingen, Department of Economics.

—— (2013). 'Gender Inequality in Social Institutions and Gendered Development Outcomes'. *World Development*, 45: 252-68.

Budlender, D. (2007). 'Gender Responsive Budgeting in Education'. In S. Aikman and E. Unterhalter (eds), *Practising Gender Equality in Education. Programme Insights.* Oxford: Oxfam.

Caputo, E., A. Lawson, and A. de Kemp (2011). 'Application of New Approach to the Evaluation of Budget Support Operations: Findings from Mali, Zambia, Tunisia. Synthesis of Main Results'. Brussels: European Commission.

Center for Global Development (2006). 'When Will We Ever Learn? Improving lives through Impact Evaluation'. Washington, DC: Center for Global Development.

Christensen, Z., D. Homer, and D. L. Nielsen (2011). 'Dodging Adverse Selection: How Donor Type and Governance Condition Aid's Effects on School Enrollment'. *World Development,* 39 (11): 2044-53.

Cronqvist, L. (2006). 'Tosmana—Tool for Small-N analysis (Version 1.3.2.)'. Trier. University of Trier: Available at: www.tosmana.net

Deiniger, K. (2003). 'Does the Cost of Schooling Affect Enrollment by the Poor? Universal Primary Education in Uganda'. *Economics of Education Review*, 22 (3): 291-305.

Demery, L. (2002). 'Gender and Public Spending: Insights from Benefit Incidence Analysis'. In UNIFEM (ed.), *Gender Budget Initiatives: Strategies, Concepts and Experiences*. New York: UNIFEM.

Dreher, A., P. Nunnenkamp, and R. Thiele (2008).'Does Aid for Education Educate Children? Evidence from Panel Data'. *The World Bank Economic Review*, 22 (2): 291-314.

EC/UN Partnership on Gender Equality for Development and Peace (2008). 'Gender Equality and Aid Effectiveness: Global Messages and Country Experiences'. Turin: EC/UN Partnership on Gender Equality for Development and Peace.

Eyben, R. (2007). 'Gender Equality and Aid Effectiveness. Challenges and Opportunities for International Practice: Experience from South East Asia'. Report prepared for a workshop organized by DFID, UNIFEM & World Bank in Bangkok (2-3 April 2007).

—— (2010). 'Subversively Accommodating Feminist Bureaucracies and Gender Mainstreaming'. *IDS Bulletin,* 41: 54-61.

Elson, D., and R. Mc Gee (1995). 'Gender Equality, Bilateral Program Assistance and Structural Adjustment: Policy and Procedures'. *World Development*, 23 (11): 1987-94.

Emmenegger, P. (2011). 'Job Security Regulations in Western Democracies: A Fuzzy Set Analysis'. *European Journal of Political Research*, 50: 336-64.

Fiss, P. (2011). Building Better Causal Theories: A Fuzzy Set Approach to Typologies in Organization Research. *Academy of Management Journal*, 54: 393-420.

Gaynor, C. (2006). 'Paris Declaration Commitments and Implications for Gender Equality and Women's Empowerment'. Paris: OECD/DAC, Network on Gender Equality.

—— (2007). 'The Paris Declaration on Aid Effectiveness and Gender Equality'. New York: United Nations, Division for the Advancement of Women.

Grofman, B., and C. Schneider (2009).'An Introduction to Crisp-set QCA, with a Comparison to Binary Logistic Regression'. *Political Research Quarterly*, 62: 662-72.

Hafner-Burton, E. M., and M. A. Pollack (2009). 'Mainstreaming Gender in the European Union: Getting the Incentives Right'. *Comparative European Politics*, 7 (1): 114-38.

Hawken, A., and G. L. Munck (2013). 'Cross-national Indices with Gender-differentiated Data: What do They Measure? How Valid are They?'. *Social Indicators Research*, 111: 801-38.

Holvoet, N. (2010). 'Gender Equality and New Aid Modalities: Is Love Really in the Air?'. *European Journal of Development Research*, 22 (1): 97-117.

Holvoet, N., and L. Inberg (2008). 'Integrating Gender-responsive Budgeting into the Aid Effectiveness Agenda. Mozambique Country Report (UNIFEM multi-country research 'GRB and Aid Effectiveness')'. Antwerp: Institute of Development Policy and Management.

—— (2011). 'Gender Mainstreaming within the Context of Changing Aid Modalities: Evidence from Tanzania'. IOB Working Paper 2011/03. Antwerp: Institute of Development Policy and Management, University of Antwerp.

—— (2012). 'Changing Aid Modalities through a Gender Lens: An International Perspective and the Case of Dutch Development Cooperation'. *Journal of International Women's Studies*, 13 (3): 1-16.

—— (*forthcoming*). 'Gender Mainstreaming in the Context of Changing Aid Modalities: Insights from two Paris Declaration Champions'. *Journal of International Development.*

IMF (2012). 'World Economic Outlook Database'. Washington DC: IMF.

Kalk, A., F. A. Paul, and E. Grabosch (2010). '"Payment for performance" in Rwanda: Does it Pay Off?' *Tropical Medicine and International Health*, 15 (2): 182-90.

Kane, E. (2004). 'Girls' Education in Africa. What Do We Know About Strategies That Work?'. Africa Region Human Development Working Paper Series. Washington, DC: World Bank.

Katz, D., and R. L. Kahn (1978). *The Social Psychology of Organizations* (Second Edition). New York: Wiley.

Killick, T. (1998). *Aid and the Political Economy of Policy Change*. London: Routledge.

King, E., and M. A. Hill (1993) (eds). *Women's Education in Developing Countries: Barriers, benefits and policies.* Baltimore, MD: John Hopkins University Press.

Lewin, K. M. (2005). 'Planning post-primary education: Taking targets to task'. *International Journal of Educational Development*, 25: 408-22.

Liverani, A., and H. E. Lundgren (2007). 'Evaluation Systems in Development Aid Agencies: An Analysis of DAC Peer Reviews 1996-2004'. *Evaluation*, 13 (2): 241-56.

Marx, A. (2010). 'Crisp-set Qualitative Comparative Analysis (csQCA) and Model Specification: Benchmark for Future csQCA Applications. *International Journal of Multiple Research Approaches*, 4: 138-58.

Marx, A., and A. Dusa (2011). Crisp-set QCA (csQCA), Contradictions and Consistency Benchmarks for Model Specification. *Methodological Innovations*, 6: 97-142.

Michaelowa, K., and A. Weber (2007). 'Aid Effectiveness in Primary, Secondary and Tertiary Education'. Background paper prepared for the 'Education for All Global Monitoring 2008'. Paris: UNESCO.

Molenaers, N., and R. Renard (2008). 'Policy Dialogue under the New Aid Approach: Which Role for Medium-sized Donors? Theoretical Reflections and Views from the Field. Discussion Paper 2008.05. Antwerp: Institute of Development Policy and Management.

Molyneux, M., and S. Razavi (2005). 'Beijing Plus 10: An Ambivalent Record on Gender Justice'. *Development and Change*, 36 (6): 984-1010.

Montinola, G. R. (2010). 'When Does Aid Conditionality Work?' *Studies in Comparative International Development*, 45: 358-82.

Moser, C., and A. Moser (2005). 'Gender Mainstreaming since Beijing: A Review of Duccesses and Limitations in International Institutions'. *Gender and Development*, 13 (2): 11-22.

OECD (nd). The Social Institutions and Gender Index (SIGI). Paris: OECD.

—— (2011). 'Findings from the Gender Equality Module of the 2011 Paris Declaration Monitoring Survey'. Paris: OECD.

OECD/DAC (2005). 'Paris Declaration on Aid Effectiveness'. Paris: OECD/DAC.

—— (2007). 'Gender Equality and Aid Delivery. What has Changed in Development Cooperation Agencies since 1999?' Paris: OECD/DAC.

—— (2008a). 'Finding the Entry Points. Gender Equality, Women's Empowerment and the Paris Declaration on Aid Effectiveness'. Issues Brief 2. Paris: OECD/DAC, DAC Network on Gender Equality.

—— (2008b). 'The DAC Gender Equality Policy Marker'. Excerpt from: Reporting Directives for the Creditor Reporting System (DCD/DAC(2007)39/FINAL). Joint Biennial Workshop of GENDERNET and IANWGE, 28-29 January 2008.

—— (2010). 'Optional Module: Gender Equality in the 2011 Survey on Monitoring the Paris Declaration'. Paris: OECD/DAC, Network on Gender Equality.

—— (2012). 'Aid in Support of Gender Equality and Women's Empowerment: Statistical Overview'. Paris: OECD DAC.

Overseas Development Institute (2012). 'From High to Low Aid: A Proposal to Classify Countries by Aid Receipt'. Background Note. London: Overseas Development Institute.

Particip (2010). 'Thematic Global Evaluation of European Commission Support to the Education Sector in Partner Countries (Including Basic and Secondary Education). Final Report Volume 1'. Freiburg: Particip.

Pollack, M. A., and E. M. Hafner-Burton (2010). 'Mainstreaming International Governance: The Environment, Gender, and IO Performance in the European Union'. *Review of International Organisations*, 5: 285-313.

Ragin, C. (1987). *The Comparative Method: Moving Beyond Qualitative and Quantitative Strategies*. Berkeley: University of California Press.

Razavi, S. (1999). 'Gendered Poverty and Well-being: Introduction'. *Development and Change*, 30: 409-33.

Razavi, S., and C. Miller (1995). 'From WID to GAD: Conceptual Shifts in the Women and Development Discourse'. UNRISD Occasional Paper Series OBP. Geneva: UNRISD.

Rihoux, B., and B. Lobe (2009). 'The Case for QCA: Adding Leverage for Thick Cross-case Comparison', In D. Byrne and C. Ragin (eds), *The Sage Handbook of Case-based Methods*. London: Sage.

Rihoux, B., and A. Marx (2013). 'Qualitative Comparative Analysis at 25: State of Play and Agenda'. *Political Research Quarterly*, 66: 167-71.

Rihoux, B., and C. Ragin (2009) (eds). *Configurational Comparative Methods: Qualitative Comparative Analysis (QCA) and Related Techniques.* Thousand Oaks: Sage.

Roberts, J. (2003).'Poverty Reduction Outcomes in Education and Health: Public Expenditure and Aid'. ODI Working Paper 210. London: Overseas Development Institute.

Rugh, A. (2000). *Starting Now: Strategies for Helping Girls Completing Primary*. Washington, DC: Sage, Academy for Educational Development.

Saith, A. (2006). 'From Universal Values to Millennium Development Goals: Lost in Translation'. *Development and Change*, 37 (6): 1167-99.

Sedelmeier, U. (2009). 'Post-accession Compliance with EU Gender Equality Legislation in Post-communist New Member States'. *European Integration Online Papers*, 2 (13): art. 23.

Schneider, C. Q., and C. Wagemann (2010). 'Standards of Good Practice in Qualitative Comparative Analysis'. *Comparative Sociology*, 9: 397-418.

Schneider, C. Q., and C. Wagemann (2012). *Strategies for Social Inquiry. Set-Theoretic Methods for the Social Sciences. A Guide to Qualitative Comparative Analysis*. Cambridge: Cambridge University Press.

Social Development Direct, Oxford Policy Management, Working Together Ltd (2008). 'Making Aid More Effective Through Gender, Rights and Inclusion: Evidence from Implementing the Paris Declaration. Analytical Summary'.

Sutherland-Addy, E. (2008). 'Gender Equity in Junior and Senior Secondary Education in Sub-Saharan Africa'. Working Paper No. 140. Washington, DC: World Bank.

Tomasevski, K. (2006) 'The State of the Right to Education Worldwide. Free or Fee: 2006 Global Report'. Copenhagen.

UNESCO (2008). 'Education for All by 2015. Will we Make it?' EFA Global Monitoring Report 2008. Paris: UNESCO.

—— (2009). 'Overcoming inequality: why Governance Matters'. EFA Global Monitoring Report 2009. Paris: UNESCO.

—— (2011). 'The Hidden Crisis: Armed Conflict and Education'. EFA Global Monitoring Report 2011.Paris: UNESCO.

—— (2012). 'Youth and Skills: Putting Education to Work'. EFA Global Monitoring Report 2012. Paris: UNESCO.

——— (2013). 'Youth and Skills: Putting Education to Work. Gender Overview'. 'EFA Global Monitoring Report 2012. Paris: UNESCO.

UNGEI (2012). 'Engendering Empowerment: Education and Equality'. New York: UNICEF.

UN Women (2010). 'Integrating Gender Responsive Budgeting into the Aid Effectiveness Agenda'. New York: UN Women.

Van Reissen, M., and M. Ussar (2005). 'Accountability Upside Down: Gender Equality in a Partnership for Poverty Eradiction'. Brussels: Eurostep and Social Watch.

Van Staveren, I. (2008). 'The Gender Bias of the Poverty Reduction Strategy Framework'. *Review of International Political Economy*, 15 (2): 289-313.

Vandemoortele J. (2002). 'Shortcuts to Public Service Provision? User Fees and Narrow Targeting'. In UNIFEM (ed.), *Gender Budget Initiatives. Strategies, Concepts and Experiences*. New York: UNIFEM.

Van der Vleuten, A., and M. Verloo (2012). 'Ranking and Benchmarking: The Political Logic of New Regulatory Instruments in the Fields of Gender Equality and Anti-corruption'. *Policy and Politics*, 40 (1): 73-88.

Weaver, C. (2008). 'The Strategic Construction of the World Bank's Gender and Development Agenda'. Paper presented at The International Studies Association Annual Meeting, 25 March, San Francisco, CA.

White, H. (2002) 'The Road to Nowhere? Results-based Management in International Co-operation: Current Issues, Food for Thought, Background Information'. Available at: euforic.org/detail_page.phtml?lang=en&page=resource_briefing_evaluat_edit3 (accessed November 2009).

Wood, B., D. Kabell, F. Sagasti, and N. Muwanga (2008). 'Evaluation of the Implementation of the Paris Declaration. Phase I. Synthesis Report.' Copenhagen.

Wood, B., J. Betts, F. Etta, J. Gayfer, D. Kabell, N. Ngwira, F. Sagasti, and M. Samaranayake (2011). 'The Evaluation of the Paris Declaration Phase II. Final Report'. Copenhagen.

World Bank (2001). 'Gender in the PRSPs: A Stocktaking'. Washington DC: World Bank, Gender and Development Group.

——— (2009). 'Abolishing School Fees in Africa. Lessons from Ethiopia, Ghana, Kenya, Malawi and Mozambique'. Washington, DC: World Bank.

World Bank and UNICEF (2009) 'Abolishing School Fees in Africa: Lessons Learned in Ethiopia, Ghana, Kenya and Mozambique'. Washington, DC/New York, World Bank/UNICEF.

Appendix 1: Overview of scores on outcome variables, aid and country context conditions

Country	AID (FINANCIAL)		CONTEXT			OUTCOME VARIABLES					
	Total aid to basic education/ primary school age child (average 2005-10) (constant 2010 US$)	Net ODA/GNI (average 2005-10) (%)	GDP/cap (average 2005-10) (PPP)	Free Education (yes/no)	SIGI (0-1) (2009)	Female NER 2005	Female NER 2010	Change in NER (2005-10)(%) (b)	Female survival rate last grade 2005	Female survival rate to last grade 2009	Change in female survival rate to last grade (2005-09)(%) (c)
Botswana	40.3	1.9	14100	no	0.08102	84	88	4.8	78	95	21.8
Burkina Faso	42.5	13.1	1141	no	0.16161	40	61	52.5	66	67	1.5
Burundi	15.8	32.8	535	yes	0.10691	58	-	-	83	61	-26.5
Ethiopia	13.7	13.7	832	yes	0.23325	66	79	19.7	59	48	-18.6
Gambia	25.3	12	1703	yes	0.17830	77	67	-13.0	-	59	-
Ghana	27.3	6.3	2385	yes	0.11269	70	84	20	-	69	-
Kenya	11.5	4.8	1564	yes	0.13704	79	83	5.1	-	-	-
Lesotho	20.3	5.9	1566	yes	(a)	89	75	-15.7	71	76	8.5
Malawi	21.2	20.3	707	yes	0.14323	97	99	2.1	36	54	50
Mali	56.5	13.3	1016	no	0.33949	45	59	31.1	70	74	5.7
Mozambique	35.8	22.4	869	yes	0.19954	74	87	17.6	39	26	-33.3
Niger	14.7	12.9	697	no	0.17559	33	57	72.7	50	67	34
Rwanda	33.3	19.7	924	yes	0.16859	75	-	-	32	39	21.9
Senegal	35.3	8.1	1818	yes	0.11041	67	78	16.4	53	61	15.1
Tanzania	16.7	13.2	1297	yes	0.11244	97	98	1.0	85	87	2.4
Uganda	10.5	13	1170	yes	0.18718	-	92	-	25	32	28
Zambia	38.7	11.4	1322	yes	0.21939	89	92	3.4	73	52	-28.8

Source: IMF (2012); Tomasevski (2006); UNESCO (2008; 2009; 2011; 2013); World Bank and UNICEF (2009); http://data.worldbank.org; http://genderindex.org/ranking_2009

(a) Lesotho has no 2009 SIGI score, as not all data was available on all SIGI sub-indicators (we based our '1' score in the dichotomous data table on the data available); (b) Change in female NER between 2005 and 2010 is calculated as (NER 2010—NER 2005)/NER 2005*100; (c) Change in female survival rate to last grade is calculated as (female survival rate 2009—female survival rate 2005)/female survival rate 2005 * 100.

Appendix 2: Dichotomous data table

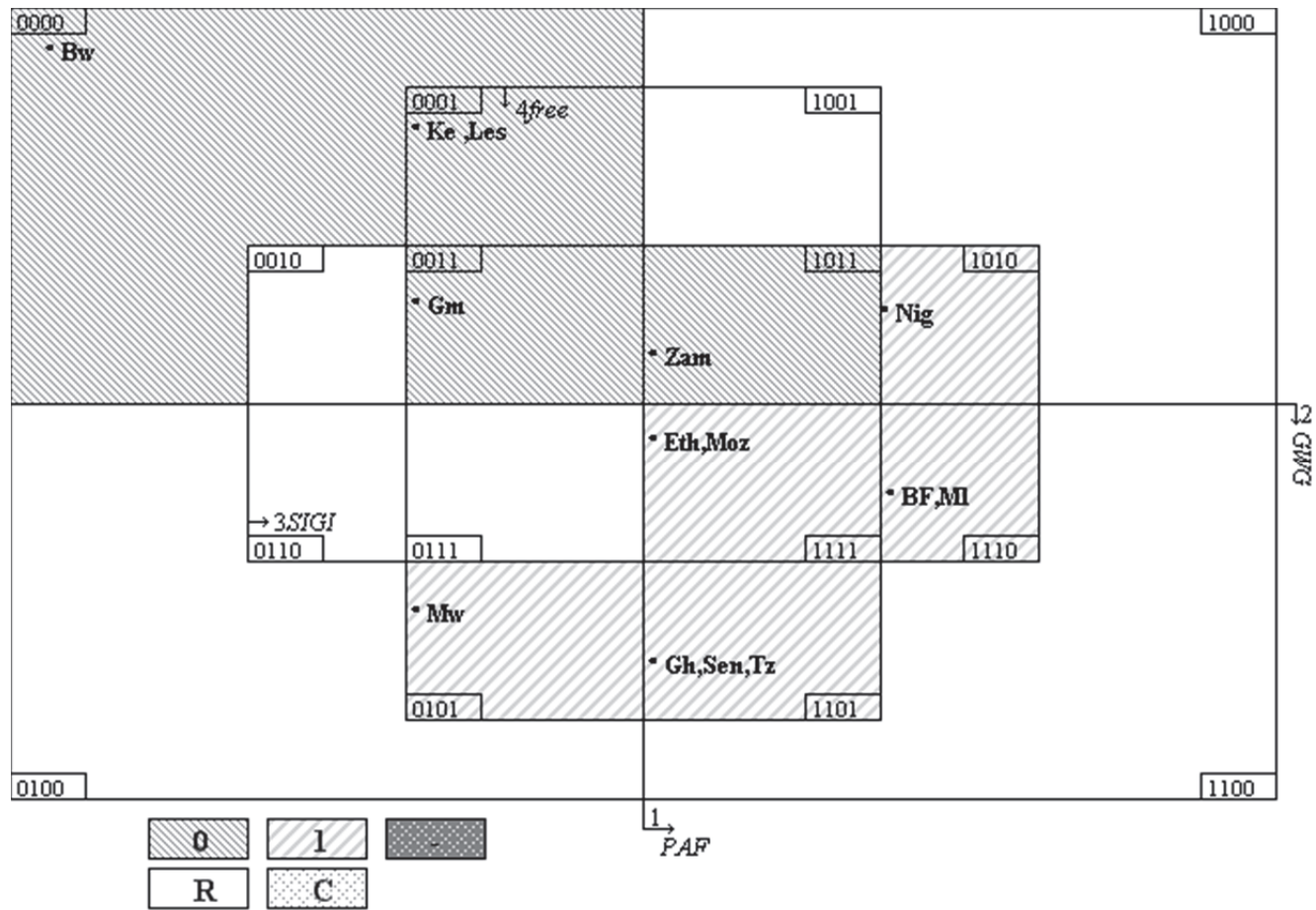| Country | AID (FINANCIAL) | | | CONTEXT | | | OUTCOME VARIABLES | |
	Total aid to basic education/ primary school age child (average 2005-10)	Net ODA/GNI (average 2005-10)	GDP/cap (average 2005-10)	Free Education	SIGI (0-1) (2009)		Change in NER (2005-10)	Change in female survival rate to last grade (2005-09)
Botswana	1	0	1	0	1		0	1
Burkina Faso	1	1	0	0	0		1	0
Burundi	1	1	0	1	1		-	0
Ethiopia	0	1	0	1	0		1	0
Gambia	1	1	1	1	0		0	-
Ghana	1	0	1	1	1		1	-
Kenya	0	0	1	1	1		0	-
Lesotho	1	0	1	1	1		0	1
Malawi	1	1	0	1	1		1	1
Mali	1	1	0	0	0		1	1
Mozambique	1	1	0	1	0		1	0
Niger	1	1	0	0	0		1	1
Rwanda	1	1	0	1	0		-	1
Senegal	1	0	1	1	1		1	1
Tanzania	1	1	1	1	1		1	0
Uganda	0	1	0	1	0		-	1
Zambia	1	1	1	1	0		0	0

Source: authors' own compilation.

Appendix 3: Venn diagram for conditions PAF, GWG, FREE, and ODA
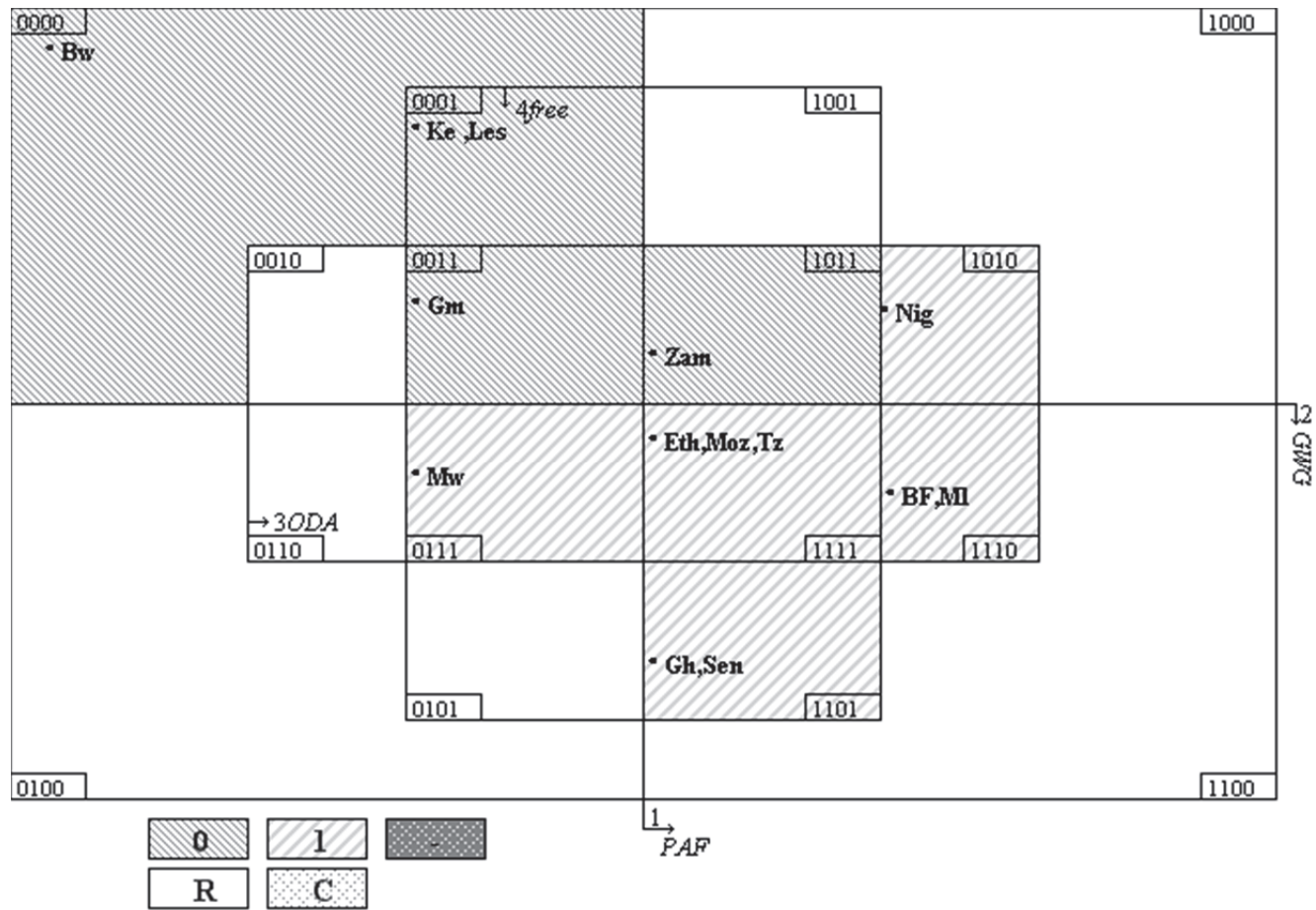
Source: output from software programme Tosmana.

Appendix 4: Venn diagram for conditions PAF, GWG, FREE and SIGI

Source: output from software programme Tosmana.