FACULTEIT ECONOMIE
EN BEDRIJFSKUNDE

TWEEKERKENSTRAAT 2
B-9000 GENT
Tel. : 32 - (0)9 – 264.34.61
Fax. : 32 - (0)9 – 264.35.92

WORKING PAPER

Is foreign aid fungible? Evidence from the education and health sectors

Nicolas Van de Sijpe

November 2010

2010/688

D/2010/7012/59

Is foreign aid fungible? Evidence from the education and health sectors*

Nicolas Van de Sijpe†

Department of Social Economics and
Study Hive for Economic Research and Public Policy Analysis (Sherppa)
Faculty of Economics and Business Administration
Ghent University

November 25, 2010

Abstract: This paper takes a fresh look at the issue of foreign aid fungibility. Unlike the bulk of existing empirical studies, I employ panel data that contain information on the specific purpose for which aid is given. This allows me to link aid given for education and health purposes to recipient public spending in these sectors. In addition, I attempt to distinguish between aid flows that are recorded on the recipient's budget and those that are off-budget, and illustrate how a failure to differentiate between on- and off-budget aid produces biased estimates of fungibility. Sector programme aid is the measure of on-budget aid, while technical cooperation serves as a proxy for off-budget aid. In both sectors, across a range of specifications, technical cooperation leads to at most a small displacement of recipient public expenditure, implying limited fungibility for this type of aid. In static fixed effects models sector programme aid shows an almost one-for-one correlation with recipient public expenditure, again suggesting low fungibility, but this effect becomes imprecise and volatile in dynamic models estimated with system GMM.

Keywords: foreign aid, fungibility, public education expenditure, public health expenditure.

JEL classification: E62, F35, H50, O23.

*I am grateful to Sudhir Anand, Channing Arndt, Steve Bond, Paolo de Renzio, Markus Eberhardt, Bernard Gauthier, Benedikt Goderis, Clare Leaver, Anita Ratcliffe, Måns Söderbom, Francis Teal, Dirk Van de Gaer, David Vines, participants at various conferences and seminars, and especially Chris Adam, Paul Collier, Shanta Devarajan, Jon Temple, Frank Windmeijer and Adrian Wood for useful comments and discussions, and to Ali Abbas, Nicolas Depetris Chauvin, Ibrahim Levent and Gerd Schwartz for help in obtaining data. I thank people at the OECD Development Assistance Committee (DAC), especially Cécile Sangaré, for patiently answering my questions about their development assistance data. I gratefully acknowledge funding from the Department of Economics, the Oxford Centre for the Analysis of Resource Rich Economies (OxCarre), the Oxford Institute for Global Economic Development (OxIGED) and St. Catherine's College, University of Oxford. This paper presents research results of the Belgian Programme on Interuniversity Poles of Attraction (contract no. P6/07) initiated by the Belgian State, Prime Minister's Office, Science Policy Programming.

†Address: Tweekerkenstraat 2, B-9000 Gent, Belgium. E-mail: Nicolas.VandeSijpe@UGent.be.

1 Introduction

In 2000 the United Nations General Assembly, consisting of 189 member countries, adopted the Millennium Declaration, laying the foundations for the Millennium Development Goals (MDGs). The eight MDGs are intended to "free our fellow men, women and children from the abject and dehumanizing conditions of extreme poverty, to which more than a billion of them are currently subjected" and to "[making] the right to development a reality for everyone and to freeing the entire human race from want" (United Nations General Assembly, 2000). Each Goal is linked to specific targets set to be attained in 2015. To achieve these targets donors acknowledge a "global partnership for development" is needed. In fact, building such a global partnership is the eighth and final goal, and it predominantly involves increasing development assistance and granting debt relief (see e.g. United Nations, 2006), at least in part to free resources for social spending.

As such, the final goal recognises the importance of external resources in attaining the MDGs. However, these external resources are unlikely to have the desired impact if they simply displace resources that recipient governments would have otherwise allocated towards meeting the MDGs. The effect of foreign aid on economic growth, poverty, and the targets set out in the MDGs may depend heavily on the recipient governments' fiscal response. One aspect of this fiscal response is the possibility that aid is fungible, i.e. that earmarked aid is used for other purposes than the one intended.

This paper endeavours to uncover to what extent earmarked education and health aid are fungible. Many studies of foreign aid fungibility are hampered by a lack of comprehensive data on the intended purpose of aid. I use the OECD's Creditor Reporting System (CRS), which contains aid disbursements disaggregated by sector or purpose, to overcome this problem. Unfortunately, however, the CRS data are incomplete. Only a fraction of the total disbursements that flow from each donor to each recipient in any given year are reported in CRS. Therefore, I propose a novel data construction method that starts from CRS and adds information from other OECD aid databases in order to come to more complete measures of education and health aid disbursements.

These data, to some extent, also allow me to split up education and health aid into an on- and off-budget component, enabling a more precise assessment of fungibility. As I illustrate in a simple analytical framework, a failure to adequately deal with the presence of off-budget aid (aid not recorded on the recipient government's budget) may have biased previous estimates of foreign aid fungibility. When donor-based measures of aid are employed in the analysis, a potentially large fraction of this aid is off-budget. Hence, even if aid is used in the targeted sector, not all of it is recorded as recipient government sectoral expenditure. This lowers the estimated marginal effect of sectoral aid on government sectoral expenditure, leading to an overestimation of the extent of fungibility. A marginal effect smaller than 1 does not necessarily mean aid is fungible, it could simply indicate that not all aid is recorded on the recipient government's budget. Other papers make use of aid data as reported by the recipient government. In this case, the effect of on-budget aid on government expenditure is estimated, while off-budget aid acts as an omitted variable. Because off-budget and on-budget aid are most likely correlated, this results in bias unless the marginal effect of off-budget aid on government spending is zero.

I use sector programme aid as a measure of on-budget aid, while technical cooperation serves as a proxy for off-budget aid. A first noteworthy finding is that technical cooperation takes up a big share of total education and health aid, suggesting that the bias from not dealing with off-budget aid in an appropriate manner is potentially large. From the analytical framework, I derive the correct null hypotheses to test for the fungibility of on- and off-budget aid.

Fixed effects estimation of a static panel data model illustrates the need to separately consider on- and off-budget aid. In both sectors, the measure of on-budget aid, sector programme aid, has an approximately one-for-one correlation with recipient public sectoral expenditure. For technical cooperation, the proxy for off-budget aid, the same result of only limited fungibility is found: its coefficient is close to and typically not significantly smaller than zero, indicating TC does not displace a recipient's own public spending in either sector. I show these results are robust to a number of specification changes and are not driven by only a handful of countries. I further employ a system GMM estimator that enables me to relax the strict exogeneity assumption implicit in the FE estimator and to consider some dynamics in the determination of public education and health expenditure. The effect of SP aid in both sectors is now estimated imprecisely and is volatile across different models, due to a lack of variation in SP aid in both sectors. Hence, no firm conclusions can be drawn with regard to SP aid. The effect of TC, however, is robust across a range of models in both sectors, and suggests that, even in the long run, TC causes at most only a small displacement of recipient public expenditure.

The next section defines fungibility and illustrates how an inappropriate treatment of off-budget aid may yield biased estimates of the degree of fungibility. It also briefly explains why aid may not be fungible. Section 3 discusses the data and the empirical model, while section 4 presents results. Section 5 concludes.

2 Fungibility and off-budget aid

2.1 Defining fungibility: the standard case

Fungibility occurs when aid is not used for the purpose intended by donors (McGillivray and Morrissey, 2004). More precisely, targeted aid is fungible if it is transformed into a pure revenue or income augmenting resource that can be spent whichever way the recipient government chooses (Khilji and Zampelli, 1994). Fungibility may arise between components of government spending (health aid could be used to finance spending on roads), in which case aid is said to be categorically fungible, or it may arise between broader fiscal aggregates (aid intended for public investment could be used for consumption purposes), which can be labelled general fungibility (McGillivray and Morrissey, 2004).

Figure 1: Graphical illustration of categorical aid fungibility

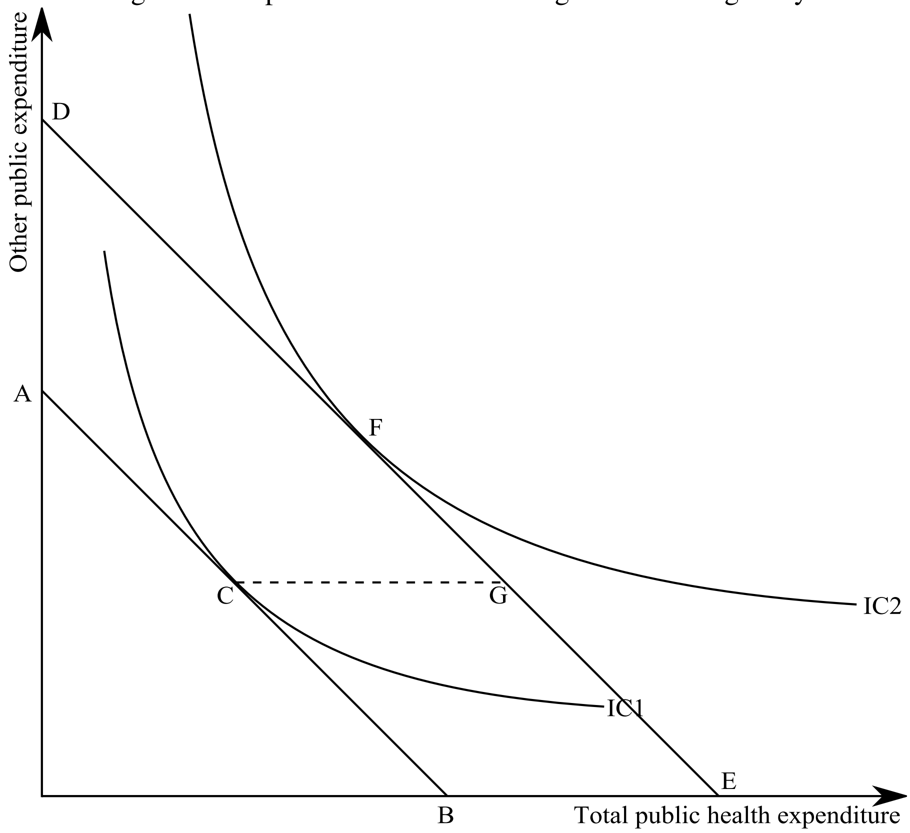


Figure 1 presents a graphical illustration of categorical aid fungibility.[1] A recipient government allocates resources between health expenditure G_H and other expenditure G_O. Given the initial budget constraint AB (with a slope of -1), government utility $V(G_H, G_O)$ is maximised at C, the point of tangency between the budget constraint and the highest achievable indifference curve (IC1). An additional amount CG of aid earmarked for health is given, shifting the budget constraint outwards to DE. Left to its own devices, the government now chooses the optimal mix of the two expenditure categories at F. Earmarked health aid is treated no differently than revenue from other sources, and is fully fungible. Fungibility would also result if the recipient government uses health aid to lower taxes or to increase the surplus or reduce the deficit. Graphically, in such cases, budget constraint DE is pushed back towards the origin, and public health expenditure again ends up being lower than it is at point G.

2.2 Taking into account off-budget aid

Some aid flows do not show up in the recipient government's budget, but may still provoke a fiscal response. I develop a simple analytical framework to illustrate how the inadequate treatment of off-budget aid has potentially biased previous fungibility estimates. From this framework, I derive the appropriate empirical tests to

[1]Similar illustrations can be found in, among others, Pack and Pack (1993), Feyzioglu, Swaroop, and Zhu (1998), and McGillivray and Morrissey (2000).

evaluate whether aid is fungible in the presence of off-budget aid.

The starting point is the following resource constraint of the recipient government:

$$G_H^{ON} + G_H^{OFF} + G_O = R + A_H^{ON} + A_H^{OFF} + A_O \quad (1)$$

where G_H^{ON} is on-budget public health spending, G_O is other public expenditure, A_H^{ON} is on-budget health aid and A_O is aid not earmarked for the health sector. R denotes unconditional resources (i.e. resources that are not earmarked for any of the expenditure categories) and is made up of domestic revenue and net borrowing. Off-budget health aid A_H^{OFF} captures aid that is not recorded on the recipient government's budget, arising from the direct provision of goods and services by donors that does not involve channelling resources through the recipient government's budget (e.g. donors building hospitals, training medical personnel…). On the expenditure side, while G_H^{ON} captures the government's health expenditure as recorded in its budget, G_H^{OFF} reflects public health spending that is not registered in the government's fiscal accounts, originating from the direct provision of goods and services by donors via off-budget health aid.[2] In other words, G_H^{OFF} reflects donor-driven spending of resources in the health sector. An important distinction between G_H^{ON} and G_H^{OFF}, therefore, is that the former is observable, whereas data on the latter are typically not available. In what follows I refer to G_H^{ON} simply as public (or government) health expenditure and to G_H^{OFF} as off-budget public health expenditure. I define *total* public health expenditure as the sum of G_H^{ON} and G_H^{OFF}.

It is logical to assume that G_H^{OFF} is financed exclusively by off-budget health aid, so it can be interpreted as the amount of off-budget health aid that remains within the health sector and is not diverted to other purposes. Public health spending undertaken from on-budget resources (e.g. taxes, on-budget aid,…) should be recorded in the budget and should therefore be part of G_H^{ON}, not G_H^{OFF}. Moreover, to the extent that the recipient government cannot trade goods and services provided directly by donors for cash or other goods and services, off-budget aid is not immediately divertible to other purposes. The mere fact that off-budget aid is excluded from budgetary records most likely reflects a lack of exclusive control of the government over these resources, so by its very nature most off-budget aid should fall into this category of aid that cannot directly be diverted to other sectors. Even if this does not hold exactly for all types of off-budget aid, in the empirical application below I focus on a specific category of off-budget sectoral aid, namely technical cooperation (e.g. the provision of experts and volunteers, the training of doctors and nurses…), for which this assumption is reasonable. Together, these two assumptions imply:

$$G_H^{OFF} = A_H^{OFF} \quad (2)$$

[2] I use the term "public" in a broad sense in this paper, to refer both to on-budget government expenditure (G_H^{ON}) and off-budget expenditure (G_H^{OFF}) originating from the direct provision of goods and services by donors via off-budget health aid.

as well as:

$$\frac{\partial G_H^{OFF}}{\partial A_H^{ON}} = 0 \tag{3}$$

In accordance with the definition of fungibility given earlier, off-budget health aid is fungible if it leads to a less than one-for-one increase in the total amount of public resources spent in the health sector (the sum of G_H^{ON} and G_H^{OFF}):

$$\frac{\partial \left(G_H^{ON} + G_H^{OFF}\right)}{\partial A_H^{OFF}} = \frac{\partial G_H^{ON}}{\partial A_H^{OFF}} + \frac{\partial G_H^{OFF}}{\partial A_H^{OFF}} < 1 \tag{4}$$

Using (2), this simplifies to:

$$\frac{\partial G_H^{ON}}{\partial A_H^{OFF}} < 0 \tag{5}$$

In other words, non-divertible off-budget health aid is fungible if it causes the government to reduce its own health spending. Full fungibility entails that the propensity to spend off-budget earmarked health aid in the health sector is no higher than the propensity to spend unconditional resources in the health sector. Algebraically, this implies:

$$\frac{\partial \left(G_H^{ON} + G_H^{OFF}\right)}{\partial A_H^{OFF}} \leqslant \frac{\partial G_H^{ON}}{\partial R} \tag{6}$$

Again using (2), off-budget health aid is fully fungible if:

$$\frac{\partial G_H^{ON}}{\partial A_H^{OFF}} \leqslant -1 + \frac{\partial G_H^{ON}}{\partial R} \tag{7}$$

Similarly, on-budget health aid is fungible if:

$$\frac{\partial \left(G_H^{ON} + G_H^{OFF}\right)}{\partial A_H^{ON}} = \frac{\partial G_H^{ON}}{\partial A_H^{ON}} + \frac{\partial G_H^{OFF}}{\partial A_H^{ON}} < 1 \tag{8}$$

After plugging in (3) this reduces to:

$$\frac{\partial G_H^{ON}}{\partial A_H^{ON}} < 1 \tag{9}$$

Full fungibility for on-budget health aid entails:

$$\frac{\partial G_H^{ON}}{\partial A_H^{ON}} \leqslant \frac{\partial G_H^{ON}}{\partial R} \tag{10}$$

Based on this simple analytical framework, I arrive at a broader, more accurate, definition of fungibility: earmarked aid is fungible if *total* public spending in the targeted sector (whether recorded on the recipient government's budget or not) increases by less than the total amount of earmarked sectoral aid (both on- and off-budget). Figure 1 can easily be re-interpreted in this light. Simply let the indifference curves reflect government preferences over total public health spending and other public expenditure, and replace the budget constraint by the resource constraint in equation (1), which includes off-budget sectoral aid.

Perhaps more importantly, this analytical framework identifies how previous studies may have produced biased fungibility estimates. Some studies rely on aid data reported by donors, either collected directly from donors or taken from databases managed by the Development Assistance Committee (DAC) of the Organisation for Economic Co-operation and Development (OECD) (e.g. McGuire, 1982, 1987; Khilji and Zampelli, 1991, 1994; Franco-Rodriguez, 2000; McGillivray and Ouattara, 2005; Osei, Morrissey, and Lloyd, 2005; Mavrotas and Ouattara, 2006; Pettersson, 2007a,b). The marginal effect of aid on recipient government expenditure is estimated and used to evaluate whether aid is fungible: the lower this marginal effect, the more fungible aid is. However, because off-budget aid, even if it is used within the targeted sector, is not counted as part of government sectoral spending, the presence of off-budget aid in the donor-based aid measure lowers the marginal effect of aid on recipient government spending, leading to an overestimation of the extent of fungibility. A marginal effect smaller than 1 does not necessarily mean aid is fungible, it could simply indicate that not all aid is recorded on the recipient government's budget. As shown in equation (5), the appropriate test to assess whether off-budget aid is fungible is to compare its marginal effect to 0, not to 1.

Other studies estimate fungibility for a single country using recipient-based aid data (e.g. Pack and Pack, 1990, 1993, 1999; Gang and Khan, 1991; Franco-Rodriguez, Morrissey, and McGillivray, 1998; Feeny, 2007).[3] In this case, the effect of on-budget aid on government expenditure is estimated, while off-budget aid acts as an omitted variable. Because off-budget and on-budget aid are most likely correlated, this results in bias unless the marginal effect of off-budget aid on government spending is zero. The sign of the bias in this case is not immediately clear, as it depends on the partial correlation between on- and off-budget aid, which could be positive or negative.

The analytical framework developed in this section therefore makes more precise McGillivray and Morrissey's (2000, p. 422) criticism that, because a large portion of aid reported by donors does not go through the recipient's public sector accounts, such aid measures "...are inappropriate for analysing fungibility." In addition, it suggests that using local aid data might not fully get around the problem of off-budget aid and may still result in biased estimates.

Off-budget aid is likely to be sizeable in many countries, so the impact of its inappropriate treatment on empirical results could be important. As far as aggregate aid is concerned, Fagernäs and Roberts (2004a) show that OECD DAC data for Uganda exceeds external financing recorded by the government by substantial margins, in some years in excess of 10% of GDP. In Zambia, the gap is as wide as 20-40% of GDP in some years (Fagernäs and Roberts, 2004b). For Senegal, Ouattara (2006) shows that OECD DAC aid during the 90s

[3] A number of studies employ both donor- and recipient-reported aid variables (e.g. Fagernäs and Schurich, 2004; Fagernäs and Roberts, 2004a,b). In addition, in a few instances it is not entirely clear whether data have been provided by recipient or donor sources. This is, for instance, the case for the sectoral loans data in Feyzioglu, Swaroop, and Zhu (1998), Swaroop, Jha, and Rajkumar (2000) and Devarajan, Rajkumar, and Swaroop (2007), which is drawn from an unpublished World Bank database.

is twice as high on average as aid reported by the local Ministry of Finance (12 vs. 6% of GDP). In Fiji and Vanuatu 70% of all aid is off-budget (Feeny, 2007), in Malawi about 40% (Fagernäs and Schurich, 2004). A recent estimate for Liberia suggests that about three quarters of aid in the fiscal year 2009-2010 is off-budget, with the percentage judged to be even higher in previous years (Republic of Liberia Ministry of Finance, 2009).

The correct way to assess whether earmarked aid is fungible is to separate on- and off-budget sectoral aid and compare the former's marginal effect on recipient sectoral spending to 1 and the latter's marginal effect to 0 (see equations (5) and (9)). The aim of this paper is to do exactly that for the education and health sectors, using a newly constructed dataset of disaggregated aid disbursements. Before turning to the empirical analysis, however, it may be worthwhile to briefly discuss some of the reasons why earmarked aid might not be fungible.

2.3 Why aid might not be fungible

As illustrated in figure 1, standard microeconomic theory predicts that fungibility arises as the natural response of a rational government to an inflow of earmarked aid. There are, however, several possible reasons that explain why aid may be less than fully fungible. Perhaps the most compelling one is donor conditionality. The earmarking of aid automatically brings with it a certain type of conditionality, namely that aid is used in the targeted sector. If the donor is able to monitor the fiscal policy choices of the recipient government and to enforce conditionality in a credible manner, full fungibility is no longer the default outcome. Adam, Andersson, Bigsten, Collier, and O'Connell (1994), for instance, set up a model in which both recipient and donor care about infrastructure and patronage spending but where the donor has a stronger relative preference for the former type of spending, as well as suffering an opportunity cost from the transfer of aid. Acting as a Stackelberg leader, the donor makes a take-it-or-leave-it offer and conditions the disbursement of earmarked aid on its use by the recipient government. The result is that, if the donor holds all bargaining power and is completely informed, it can extract a more than one-for-one increase in sectoral spending from the recipient government, so that earmarked aid is not fungible.[4]

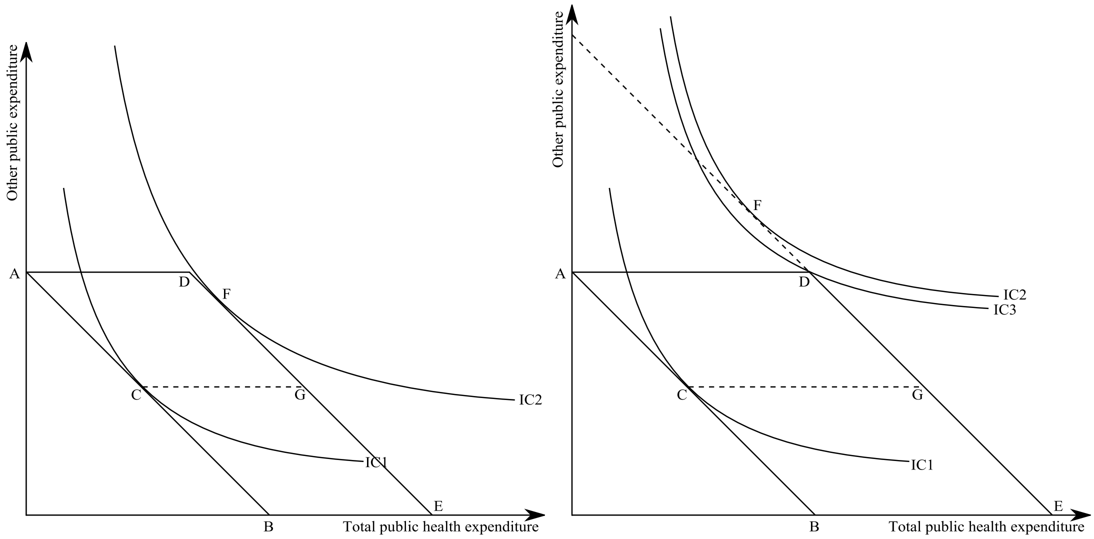
A lack of information on the recipient government's part may also reduce the degree of fungibility. McGillivray and Morrissey (2001) argue that, even if policymakers in the recipient country intend earmarked aid to be fully fungible, fungibility might be reduced due to perception errors of implementing officials ("aid illusion"). For instance, if earmarked aid is given in kind and implementing officials wrongly believe the price of the targeted good falls as a result, spending in the sector could increase by more than intended. Misperceptions may also arise if earmarked aid is given in the form of a matching grant. A matching grant effectively subsidises the purchase of a good up to a certain threshold. Graphically, in figure 1, the budget constraint would be kinked

[4] This echoes a result in Azam and Laffont (2003), who find that, under complete information, the consumption of the poor rises more than one-for-one with aid that is made conditional on the poor's consumption. For a formal analysis of a principal-agent model of aid fungibility and an extension to an incomplete information set-up, see Van de Sijpe (2010).

at the threshold level, with the part to the left of the kink being flatter: for every dollar reallocated from other expenditure to health expenditure the donor makes an additional contribution, until the agreed health expenditure threshold is reached. If the threshold is overestimated by implementing officials, fungibility might again be lower than what was intended by policy officials.

Incomplete information may contribute in particular to a reduction in the fungibility of off-budget aid. If governments in aid-receiving countries are not aware of the extent to which donors are directly providing goods and services in a sector via off-budget aid, they may not realise that total public expenditure in the sector is higher than what they consider as optimal, and, as a result, may neglect to reduce their own expenditure in the sector in the wake of an inflow of off-budget aid.

Figure 2: Graphical illustration of aid fungibility with a kinked budget constraint



(a) Fungibility not reduced (b) Fungibility reduced

There is a final reason to expect less than full fungibility that is specific to off-budget aid. The presence of off-budget health aid that cannot directly be diverted to other sectors determines a lower bound for the amount of total public health spending (the sum of on- and off-budget spending). Because of this lower bound, the budget constraint after the transfer of earmarked aid, ADE, is kinked, as shown in figure 2.[5] Despite the fact that it cannot directly be diverted to other sectors, in figure 2(a) off-budget health aid is still fully fungible, as the recipient government reduces its own sectoral spending in response to the inflow of aid. However, fungibility is reduced if the government's optimal amount of total public health expenditure is exceeded by the amount of non-divertible off-budget health aid. This is the case if the kink in the budget constraint (point D) lies to the South-East of F, as in figure 2(b).

While, at first blush, one would perhaps expect this to be a rare occurrence, it may be more relevant if we think of the government as separately targeting optimal amounts of various types of health goods that cannot easily substitute for each other, rather than one aggregate health good. In that case it would be more likely

[5]For simplicity, in the figure I assume all health aid is off-budget.

that non-divertible off-budget health aid directed towards one or several of these more specific health goods (hospitals, syringes, health technical cooperation. . .) exceeds the government's preferred expenditure on that good, so that the fungibility of earmarked health aid as a whole is brought down (Gramlich, 1977, makes exactly this point in the context of intergovernmental grants).

Ultimately then, the extent to which earmarked aid is fungible needs to be determined empirically, which I take up in the remainder of this paper.

3 Data and empirical model

3.1 Sectoral aid data

Knowing the intended purpose of aid is crucial to accurately estimate the degree of fungibility. The use of sectorally disaggregated aid in this paper therefore constitutes a marked improvement on previous studies that lack complete information on the purpose for which aid is given. Fiscal response models (FRMs) typically focus on the effect of aggregate aid on the recipient's budget, judging aid to be fungible if it is diverted away from public investment or developmental expenditure (see e.g. Heller, 1975; Franco-Rodriguez, Morrissey, and McGillivray, 1998; Feeny, 2007). Early fungibility studies, such as McGuire (1982, 1987) and Khilji and Zampelli (1991, 1994), distinguish between military and economic aid and evaluate how these affect public military and non-military expenditure. Other studies (Feyzioglu, Swaroop, and Zhu, 1998; Swaroop, Jha, and Rajkumar, 2000; Devarajan, Rajkumar, and Swaroop, 2007) further attempt to break aid down to the sectoral level but are only able to disaggregate concessionary loans, so the omission of sectoral grants may influence their results. In this literature, Pack and Pack (1990, 1993, 1999) are the only studies able to employ a comprehensive sectoral disaggregation of foreign aid, by virtue of focusing on countries whose recipient governments report both public expenditure and aid received in a disaggregated form.

In addition, a few recent studies (Chatterjee, Giuliano, and Kaya, 2007; Pettersson, 2007a,b) use sectorally disaggregated aid data from the OECD's Creditor Reporting System (CRS), described in OECD (2002), to study fungibility.[6] The CRS database allows one to disaggregate foreign aid along a number of dimensions, most importantly the sector or purpose of aid. These fungibility studies are part of a broader, fast expanding, literature that uses disaggregated aid data from CRS to examine a range of issues, including the effect of education and health aid on outcomes in these sectors (Michaelowa and Weber, 2006; Mishra and Newhouse, 2007; Wolf, 2007; Dreher, Nunnenkamp, and Thiele, 2008; Williamson, 2008), the responsiveness of sectoral aid allocations to MDG-related indicators of need (Kasuga, 2007; Thiele, Nunnenkamp, and Dreher, 2007), and the effect of various aid categories on economic growth (Mavrotas, 2002a; Clemens, Radelet, and Bhavnani, 2004; Asiedu

[6] I describe the OECD's aid databases as they were at the time I started to construct the sectoral aid data (December 2006). Since then, the CRS and DAC Directives have been updated and the databases have undergone some minor changes (see OECD, 2007a,b).

and Nandwa, 2007). A number of FRMs (Mavrotas, 2002b, 2005; Mavrotas and Ouattara, 2006; Ouattara, 2007) also make use of CRS to investigate the effects of various aid modalities (project aid, programme aid, technical assistance, and food aid) on the recipient government's budget.

While CRS is a very useful database to study topics that require disaggregated aid data, it is not without faults. Firstly, the aid data in CRS are incomplete. Only a fraction of the total disbursements that flow from each donor to each recipient in any given year are reported. Coverage becomes weaker as one goes further back in time. Secondly, while commitments are available since 1973, disbursements are only available from 1990 onwards. As a result, many existing papers make use of sectoral commitments, even if disbursements are the more relevant quantity.

A few studies (e.g. Mavrotas, 2002a,b, 2005; Michaelowa and Weber, 2006; Pettersson, 2007a,b) attempt to get around these problems with the help of data from OECD DAC table 2a, described in OECD (2000a). DAC2a contains *complete* aggregate aid disbursements, but no sectoral disaggregation. These studies estimate sectoral disbursements for each recipient and each year ($\widehat{d}^s_{RY}$) by calculating the share of each sector s in total CRS commitments, and then multiplying these shares with aggregate disbursements from DAC2a ($DAC2a^{agg}_{RY}$):[7]

$$\widehat{d}^s_{RY} = DAC2a^{agg}_{RY}\left(\frac{CRS^{s,comm}_{RY}}{CRS^{agg,comm}_{RY}}\right) \tag{11}$$

for $s = 1, \ldots, S$. This strategy yields sectoral aid disbursements even for those years where only commitments are available in CRS. Moreover, because $DAC2a^{agg}_{RY}$ is complete, it corrects for the incomplete nature of the CRS data in a simple way.

The key assumption underlying this method is that the sectoral distribution of incomplete CRS commitments is a good guide to the actual distribution of total disbursements across sectors. There are, however, several reasons why this assumption may not hold. A donor's propensity to report disaggregated aid to the CRS database may vary by sector. Related to this, donors that report a good deal of their aid to CRS might have different sectoral preferences than donors that largely fail to report disaggregated aid. In both cases, from a recipient's perspective, the incomplete CRS data might be a poor guide to the true sectoral allocation of aid received. In addition, the link between commitments in time t and disbursements in the same period might differ across sectors. For these reasons, simply scaling sectoral CRS commitments so that their sum matches aggregate DAC2a disbursements could yield highly imperfect measures of sectoral disbursements. This is especially the case if CRS coverage is low, so that the sectoral distribution of CRS commitments that is used to allocate aggregate DAC2a disbursements across sectors is based on only a small subset of the total aid committed to a recipient.

[7] RY stands for recipient-year, agg indicates aggregate aid and $comm$ denotes commitments. No superscript is used for disbursements.

To address these problems, I first of all restrict the analysis to the period 1990-2004, for which CRS disbursements are available. More importantly, in order to construct more complete data on earmarked education and health aid disbursements for this period, I propose a more elaborate data construction method that takes into account additional information available in DAC table 2a and DAC table 5. As the method is described in detail in the appendix, I provide only a brief summary here.

I start with gross CRS disbursements in a recipient-donor-year (RDY) format, i.e. showing for each year how much aid is transferred from each donor to each recipient. These aggregate and sectoral disbursements are referred to as CRS^{agg}_{RDY} and CRS^{s}_{RDY} (for $s = 1, \ldots, S$), respectively. In addition, I use data on aggregate disbursements from DAC2a, again in a RDY format ($DAC2a^{agg}_{RDY}$). $DAC2a^{agg}_{RDY}$ should be complete but cannot be decomposed by sector. For each recipient-donor-year observation the amount of aid missing from CRS is calculated as the difference between DAC2a and CRS disbursements:

$$RES^{agg}_{RDY} = DAC2a^{agg}_{RDY} - CRS^{agg}_{RDY} \tag{12}$$

The aim is to allocate this total residual (RES^{agg}_{RDY}) across sectors, thereby generating sectoral residuals that can be added to the CRS sectoral disbursements to make up for the incomplete nature of the latter.

To accomplish this, data from one more table is needed. DAC table 5 comprises aggregate aid and its sectoral distribution but only by donor. While this means the data are not available from a recipient perspective, the advantage of DAC5 is that it should contain more complete information than CRS. From DAC5 I obtain data on aggregate aid and its sectoral allocation for each donor and each year ($DAC5^{agg}_{DY}$ and $DAC5^{s}_{DY}$, respectively).[8] I sum the above CRS data over all recipients to get it in the same donor-year (DY) format. For each sector I can then calculate the amount of sectoral aid missing from CRS in each donor-year. In addition to the total residual already calculated, this yields a residual for each sector (RES^{s}_{DY}). As a result, for each donor-year and sector I can work out the share of the sectoral residual in the total residual:

$$SHRES^{s}_{DY} = \frac{RES^{s}_{DY}}{\sum_{s=1}^{S} RES^{s}_{DY}} \tag{13}$$

This donor- and year-specific allocation of the total residual across sectors is then applied to the total residual in the original recipient-donor-year format:

$$\widehat{RES}^{s}_{RDY} = SHRES^{s}_{DY} RES^{agg}_{RDY} \tag{14}$$

That is, I apply the sectoral residual shares of a given donor-year to the total residuals of all recipients to which

[8]Unfortunately, the data in DAC5 is a mix of disbursements and commitments. To account for this I scale the DAC5 data so that the sum of the sectoral aid variables matches aggregate disbursements from DAC2a in every donor-year, as explained in more detail in the appendix.

the donor gives aid in that year that is not fully accounted for in CRS. In other words, I assume the sectoral allocation of a donor's total residual is the same for all recipients with which this donor has a residual. This yields sectoral residual variables in a recipient-donor-year format ($\widehat{RES}^{s}_{RDY}$), which are added to CRS sectoral disbursements to create more complete measures of sectoral aid:

$$\widetilde{CRS}^{s}_{RDY} = CRS^{s}_{RDY} + \widehat{RES}^{s}_{RDY} \tag{15}$$

Summing across donors brings the sectoral disbursements in the required recipient-year format:

$$\widetilde{CRS}^{s}_{RY} = \sum_{D} \widetilde{CRS}^{s}_{RDY} \tag{16}$$

For some donors insufficient information is available in DAC5 to allocate the total residual across sectors, so for some observations the constructed sectoral aid variables still do not reflect the total amount of aid received. Therefore, as a final step, I scale the sectoral disbursements so that their sum matches a plausible measure of aggregate disbursements received ($DISB_{RY}$, see the appendix for details):

$$\widehat{CRS}^{s}_{RY} = DISB_{RY} \left(\frac{\widetilde{CRS}^{s}_{RY}}{\sum_{s=1}^{S} \widetilde{CRS}^{s}_{RY}} \right) \tag{17}$$

Aid disbursements are constructed for the following sectors: education (DAC5 sector code 110), health (120), commodity aid/general programme assistance (500), action relating to debt (600), donor administrative costs (910), support to NGOs (920) and other sectors (the sum of all remaining sector codes).[9] In addition, following a similar procedure as the one described above at the sector level, data that partition education and health disbursements into four prefix codes or aid types are constructed. These prefix codes are investment projects (IP), sector programme (SP) aid, technical cooperation (TC), and other (no mark) (ONM). Definitions and details on the construction of the prefix codes can be found in the appendix. As explained in more detail below, the prefix codes are useful because, to some extent, they allow separation of on- and off-budget aid flows, enabling a test of fungibility in line with the analytical framework discussed in section 2.

The strategy pursued here to construct sectoral aid disbursements tries to take into account that donors that report only a small part of their aid to CRS might allocate aid across sectors in a different way than donors that report a larger part of their aid. Similarly, it takes into account that, for a given donor, the sectoral allocation of unreported aid might be different from that of the reported portion. Sectoral aid disbursements are constructed in such a way as to assure the distribution of aggregate aid across sectors for each donor-year closely follows the sectoral allocation in DAC5, which contains complete disaggregated aid data. After this, the main assumption

[9] In CRS, the sector is recorded using a 5-digit purpose code, the first 3 digits of which refer to the corresponding sector in DAC5 (see OECD, 2002, Annex 5, pp. 87-106). It is these 3 digits I focus on here to delineate sectors.

is that the sectoral allocation of the total residual in donor-year format applies equally to each recipient that receives aid from the donor in that year that is not accounted for in CRS. While, for a given donor, the sectoral allocation of the total residual might differ across recipients, information on this is not available, so this is the best approximation that can be made.

In the final step of the data construction I scale the sectoral aid variables so that their sum matches a measure of aggregate aid received ($DISB_{RY}$), similar to what has been done in previous studies (recall equation (11)). However, because the sectoral disbursements before scaling are based on more extensive information than in previous studies, they are more likely to be a useful guide to the true sectoral allocation of total disbursements and the scaling should therefore be less problematic. Table 11 on p. 48 in the appendix shows summary statistics for the scaling that takes place in the last step.[10] $scaling$ is calculated as the ratio of the sum of the constructed sectoral disbursements (before scaling) to $DISB_{RY}$:

$$scaling = \frac{\sum_{s=1}^{S} \widetilde{CRS}_{RY}^{s}}{DISB_{RY}} \tag{18}$$

This is compared to the scaling that would take place if I simply scale sectoral CRS disbursements as some existing studies have done:

$$scaling_{CRS} = \frac{\sum_{s=1}^{S} CRS_{RY}^{s}}{DISB_{RY}} \tag{19}$$

As can be seen from table 11, the difference between $scaling$ and $scaling_{CRS}$ is large. On average, the constructed disbursements before scaling make up more than 76% of aggregate, complete disbursements, whereas for CRS disbursements this is only 31.9%. This difference reflects the information added to the sectoral CRS disbursements by the data construction method developed here. For the majority of observations the scaling performed in the final step of the data construction is limited in magnitude, and a lot smaller than if CRS sectoral disbursements are scaled without any adjustment. For instance, for more than three quarters of the observations CRS disbursements constitute less than half of aggregate aid. For the constructed sectoral disbursements this is the case for less than 10% of observations. This makes it more likely that the sectoral allocation of the aid data before scaling is a reasonable reflection of the actual sectoral allocation one would find if data were complete. This is again the best that can be done with the available data, and not scaling the sectoral disbursements runs the risk of underestimating the amount of aid received.[11]

[10] As the empirical analysis later is restricted to low and middle income countries, I exclude high income countries to calculate the figures in table 11.

[11] Since constructing the data for this paper, two new disaggregated aid datasets have become available. Ravishankar, Gubbins, Cooley, Leach-Kemon, Michaud, Jamison, and Murray (2009) construct data on health aid by estimating disbursements on the basis of the less incomplete CRS commitments and by adding data, obtained from separate reports, for a number of NGOs and multilateral and private donors. These data are used by Lu, Schneider, Gubbins, Leach-Kemon, Jamison, and Murray (2010) to estimate the fungibility of health aid. One downside is that a large part of the data cannot be allocated by recipient country. Lu, Schneider, Gubbins, Leach-Kemon, Jamison, and Murray (2010, p. 1379) state that the health aid that could be traced directly to recipient countries represents 21% of all health aid in 1995 and 30% in 2006. The remainder is made up by "resources given to organisations, or activities that were regional or global or that could not be traced to specific countries". In addition, it is not immediately clear how one would further split

3.2 Empirical model and other data

The sectoral aid data allow me to estimate models that relate recipient government sectoral expenditure to aid earmarked for that sector:

$$SSP_{it} = \beta SAID_{it} + \boldsymbol{\gamma} \boldsymbol{A}_{it} + \boldsymbol{\delta} \boldsymbol{X}_{it} + \lambda_t + \eta_i + \epsilon_{it} \tag{20}$$

for $i = 1, \ldots, N$ and $t = 1, \ldots, T$. SSP_{it} denotes recipient government spending on education or health, while $SAID_{it}$ are disbursements earmarked for the same sector. β is the main parameter of interest. $\boldsymbol{A}_{it}$ and $\boldsymbol{X}_{it}$ are column vectors containing other aid variables and control variables, respectively, as described in more detail below, and $\boldsymbol{\gamma}$ and $\boldsymbol{\delta}$ are row vectors of parameters. λ_t is a set of year dummies, η_i captures country-specific time-invariant effects and ϵ_{it} is the transient error. Aid and spending variables are expressed as a percentage of GDP.[12] I restrict the analysis to low and middle income aid recipients, so high income countries (2005 GNI per capita of 10726 US$ or more, following World Bank, 2006c) are dropped from the sample. I start from a static panel data model as this stays close to what has been done by those cross-country fungibility studies that have some information on the intended purpose of aid, in particular Feyzioglu, Swaroop, and Zhu (1998) and Devarajan, Rajkumar, and Swaroop (2007), and therefore allows for an easier comparison of results. Later in the paper I also estimate more general models that allow for some dynamics.

I focus on education and health for a number of reasons. Firstly, education and health play a prominent role in the Millennium Development Goals (MDGs). Apart from their importance in eradicating extreme poverty and hunger, the first goal, five other goals explicitly set targets related to education and health.[13] This suggests donors have some preference for education and health spending and, as a result, should care about the extent of fungibility in these sectors. Secondly, as partly evidenced by their prominent role in the MDGs, there is a fairly widespread belief that better education and health have immediate consequences for human welfare, as well as playing an important role in spurring development and alleviating poverty. This suggests the fungibility of aid directed towards these sectors may matter for the welfare of the population in recipient countries and have a bearing on overall aid effectiveness. Thirdly, these are rather clearly defined areas of spending, which should increase the definitional overlap between sectoral aid and sectoral spending.

Public education and health expenditure as a percentage of GDP are staff estimates from the IMF's Fiscal Affairs Department (FAD).[14] The latest update, after which collection ceased, is October 2004, so the last year

up health aid into an on- and off-budget component in these data. A second recent dataset, AidData (http://www.aiddata.org), attempts to construct a fuller disaggregation of aggregate aid into all its constituent parts along a number of dimensions, but focuses almost exclusively on commitments.

[12] Current US$ GDP from World Bank (2006c) is used to express sectoral aid disbursements as a percentage of GDP.

[13] These are achieving universal primary education, promoting gender equality and empowering women (primarily measured by gender disparity in primary and secondary education), reducing child mortality, improving maternal health, and combatting HIV/AIDS, malaria and other diseases (see e.g. United Nations, 2006).

[14] These data are not publicly available, though it has been used in a variety of publications (Gupta, Clements, and Tiongson, 1998;

with data is 2003. The data are taken from IMF country documents and checked by the desk economists for each country for verification and reconciliation (Baqir, 2002). The main advantage over other datasets (International Monetary Fund, 2006; World Bank, 2006a,c) is greater coverage. Also, while the level of government (central or general, where the latter also includes state and local government spending) differs across countries, it is fixed over time, so average differences in public education and health spending between countries due to differences in the government level on which reporting is based can be picked up by fixed effects (Baqir, 2002).[15]

$\boldsymbol{A}_{it}$ includes commodity aid/general programme assistance (henceforth called general aid) and support to NGOs. If targeted towards education and health, support to NGOs may have an effect on the recipient government's own spending in these sectors, while general aid may partly finance education and health spending. On the other hand, parts of general aid may be linked to structural adjustment programmes and may therefore be conditional on lowering public spending. The third and final variable in $\boldsymbol{A}_{it}$ is other non-education or other non-health aid. In the equation for public education spending other non-education aid includes health aid. Similarly, in the equation for public health spending, other non-health aid contains education aid.

The remaining two aid variables for which I have constructed data, action relating to debt and donor administrative costs, are not included in the regression model. Although donor administrative costs can be allocated by recipient they should have no bearing on the recipient government's fiscal policy decisions. Debt relief may affect public social spending and at the same time be correlated with the amount of education and health aid but it is not adequately captured by action relating to debt, which consists of debt forgiveness, debt rescheduling, and other action on debt (such as service payments to third parties, debt conversions, and debt buybacks) (OECD, 2000b). The debt forgiveness component measures the face value of total debt forgiven in a year rather than its present value (PV). Because the average concessionality of debt varies strongly across countries this may be misleading (Depetris Chauvin and Kraay, 2005). For most types of debt rescheduling the reduction in debt service in a given year due to present and past rescheduling is recorded. Again, this fails to capture the PV of current and future reductions in debt service due to debt rescheduling in the current year.[16]

For these reasons, I omit action relating to debt as a regressor, and instead control for the PV of public and publicly guaranteed long-term external debt, as well as public and publicly guaranteed long-term external debt service. These variables should pick up most of the effect of debt relief on social spending. The PV of debt is taken from Dikhanov (2004), updated through to 2004.[17] It has the additional advantage over action relating to

Gupta, Dicks-Mireaux, Khemani, McDonald, and Verhoeven, 2000; Baqir, 2002; Thomas, 2006; Lora and Olivera, 2007). I am very grateful to Gerd Schwartz for sharing this data and to Ali Abbas for help in obtaining it.

[15]For Fiji the observation in 1998 for both sectors is about ten times smaller than that in the surrounding years, most likely due to a typographical error. Public education expenditure, for instance, is 0.572% of GDP in 1998, whereas it hovers between 5.19 and 6.37% of GDP in all other years from 1993 to 2002. Hence, I change this value to 5.72. Similarly, I change the public health expenditure value for 1998 from 0.253 to 2.53% of GDP.

[16]Only for Paris Club concessional debt reorganisations is the net present value reduction in debt achieved by current rescheduling recorded (OECD, 2000b, p. 17).

[17]I am very grateful to Ibrahim Levent for sending me the updated data (received December 2006) as well as the Dikhanov paper.

debt that it takes into account (reductions in) debt owed to creditors that are not included in the DAC, as it is based on debtor-based data from the World Bank's Debtor Reporting System. The source for debt service is the Global Development Finance database (World Bank, 2006b). I again use current US$ GDP from World Bank (2006c) to express both variables as a percentage of GDP.

Table 1: Summary statistics

Variable	Mean	Std. Dev.	Min.	Max.
Education sector: 1082 observations				
Public education expenditure	4.02	1.92	0.38	13.61
Education aid	1.13	1.45	0.01	14.19
Education IP	0.13	0.23	0	3.6
Education SP	0.04	0.09	0	0.95
Education TC	0.81	1.1	0	10.85
Education ONM	0.16	0.34	0	5.83
General aid	1.2	1.92	0	22.78
Support to NGOs	0.13	0.24	0	3.02
Other non-education aid	5.84	6.78	0.01	62.84
Real GDP per capita	3.63	2.98	0.47	17.96
Real GDP per capita growth	1.6	5.46	-30.28	49.86
Urbanisation	42.4	20.36	6.3	91.56
Trade	78.11	41.06	10.83	280.36
PV debt	52.15	60.07	0.09	892.12
Public debt service	4.02	3.47	0	35.24
Health sector: 1087 observations				
Public health expenditure	1.96	1.25	0.17	7.44
Health aid	0.44	0.54	0	3.63
Health IP	0.11	0.18	0	1.69
Health SP	0.05	0.1	0	1.75
Health TC	0.18	0.23	0	1.91
Health ONM	0.1	0.18	0	1.46
General aid	1.21	1.97	0	22.78
Support to NGOs	0.13	0.24	0	3.02
Other non-health aid	6.56	7.5	0.02	66.11
Real GDP per capita	3.64	2.98	0.47	17.96
Real GDP per capita growth	1.58	5.4	-30.28	28.5
Urbanisation	42.24	20.4	6.3	91.56
Trade	77.8	41.2	10.83	280.36
PV debt	51.12	59.14	0.09	892.12
Public debt service	3.91	3.24	0	35.24

Note: All variables as % of GDP except real GDP per capita (thousands of constant 2000 international dollars) and its growth rate, and urbanisation (urban population, % of total).

Other control variables included in $\boldsymbol{X}_{it}$ (all taken from World Bank, 2006c) are real GDP per capita (thousands of constant 2000 international dollars) and its growth rate, urbanisation (urban population, % of total) and trade (sum of imports and exports, % of GDP). Since aid expressed as a % of GDP is very likely to be correlated with GDP (per capita), excluding the latter from the equation may induce a spurious relationship between sectoral aid and expenditure. The growth variable is included to capture the reaction of public sectoral expenditure to short-run shocks in GDP per capita, which may also influence the allocation of aid. If government education and health expenditure do not immediately adjust to a higher (lower) level in the wake of a positive (negative) growth shock, a negative coefficient is expected. The effect of trade on public education and health expenditure is a priori ambiguous (see e.g. Rodrik, 1998; Kaufman and Segura-Ubiergo, 2001; Dreher, 2006; Dreher, Sturm, and Ursprung, 2006). Greater openness may erode a government's capacity to finance expenditure as tax bases become more mobile. However, openness to trade may also raise the demand for social spending to insure for increased external risk and to redistribute the gains from trade, and public education and health expenditure may play a role in this. Urbanisation, as well, could have a positive or negative effect. On the one hand, some services should be easier to administer in a more urbanised society (Hepp, 2005) and urbanisation may create more opportunities for economies of scale. On the other hand, lower transportation costs may increase the demand for education and health services (Hepp, 2005) and – mainly relevant for health spending – the risk of contagion may be higher in cities (Gerdtham and Jönsson, 2000).

Table 1 shows summary statistics for the education and health regression samples. Despite the better coverage of the FAD data the availability of observations for public education and health expenditure is still the main constraint on the sample size. Education aid makes up about 28% of public spending in the education sector, while health aid stands at about 22% of public health spending. A little bit less than a fifth of aid (excluding actions relating to debt and donor administrative costs) is targeted towards the education and health sectors. In both sectors TC is the dominant form of aid. Average SP aid in both sectors is very small, reflecting the fact that for many country-years education and health sector programme aid are all but zero.

3.3 Hypothesis tests for no and full fungibility

As I have argued in section 2.2, the presence of off-budget aid in the donor-based measure of sectoral aid ($SAID_{it}$) pulls down the coefficient estimate of β, thereby overstating the true extent of fungibility. To arrive at a more accurate assessment of fungibility it is therefore necessary to distinguish between on- and off-budget sectoral aid. Consequently, I also estimate models that partition education and health disbursements into four aid types or prefix codes:

$$\begin{aligned} SSP_{it} = {} & \beta_{IP} SAIDIP_{it} + \beta_{SP} SAIDSP_{it} + \beta_{TC} SAIDTC_{it} \\ & + \beta_{ONM} SAIDONM_{it} + \boldsymbol{\gamma} \boldsymbol{A}_{it} + \boldsymbol{\delta} \boldsymbol{X}_{it} + \lambda_t + \eta_i + \epsilon_{it} \end{aligned} \tag{21}$$

where IP stands for investment projects, SP for sector programme aid, TC for technical cooperation, and ONM for other (no mark) aid.

This further disaggregation allows a more precise test of whether education and health aid are fungible.

Sector programme aid should for the most part be on-budget, as by definition programme aid involves a government to government transfer of resources. Technical cooperation, on the other hand, should be a good proxy for off-budget aid. The cost of providing training and scholarships in donor countries, remunerating experts and consultants, and financing equipment and administrative costs associated with TC mostly involve a direct payment from the donor government rather than a transfer of money to the recipient government. Fagernäs and Roberts (2004a,b), Feeny (2007) and IDD and Associates (2006) all attribute discrepancies between donor and recipient reports of aid in the countries they are studying at least in part to the omission of technical assistance from recipient governments' budgets, while Baser and Morgan (2001) state more explicitly that, for the six African countries they investigate, TC is off-budget. The summary statistics in table 1 further reinforce the notion that the bias from not dealing with off-budget aid in an adequate manner is potentially large: education aid is more than 70% TC, while in the health sector around 40% of aid is TC. This dominant role of TC in health aid and especially education aid is confirmed in the CRS directives (OECD, 2002, p. 26).

The extent to which investment projects and other (no mark) aid are reported in recipient government budgets is more uncertain, so the estimates for β_{IP} and β_{ONM} are not very informative to gauge the degree of fungibility. However, as measures of on- and off-budget sectoral aid, respectively, for $SAIDSP_{it}$ and $SAIDTC_{it}$ it is possible to test the null hypothesis of no fungibility and the null of full fungibility in line with the analysis in section 2, as shown in table 2.

Table 2: Null hypotheses for no and full fungibility with on- and off-budget aid

Theoretical null hypotheses:	No fungibility	Full fungibility
Aid on-budget (SP)	$\beta_{SP} \geqslant 1$	$\beta_{SP} \leqslant \frac{\partial SSP_{it}}{\partial R_{it}}$
Aid off-budget (TC)	$\beta_{TC} \geqslant 0$	$\beta_{TC} \leqslant \frac{\partial SSP_{it}}{\partial R_{it}} - 1$
Implemented null hypotheses:	No fungibility	Full fungibility
Aid on-budget (SP)	$\beta_{SP} \geqslant 1$	$\beta_{SP} \leqslant 0$
Aid off-budget (TC)	$\beta_{TC} \geqslant 0$	$\beta_{TC} \leqslant -1$

Carrying out the full fungibility hypothesis tests requires knowledge of the marginal effect of unconditional resources R (typically measured as government expenditure net of aid) on public sectoral expenditure, which could be obtained by following the two-stage procedure outlined in Devarajan, Rajkumar, and Swaroop (2007).[18] The fiscal policy data I received from the IMF's FAD, however, do not contain total government expenditure, revenue or borrowing, and because data availability for these variables in the IMF's Government Finance Statistics or in the World Bank's World Development Indicators is much more limited, a large fraction of the sample would be lost by following this procedure. Instead, I set $\frac{\partial SSP_{it}}{\partial R_{it}} = 0$ to carry out the full fungibility hypothesis tests so that the implemented tests become those shown in the bottom half of table 2.

[18] Essentially, this procedure entails including the residual from a regression of R on the right hand side variables in equation (21) as an explanatory variable in the model. Since this residual is, by construction, orthogonal to the other right hand side variables, its inclusion does not influence the sectoral aid coefficient, which still captures the full effect of earmarked aid on public sectoral spending. It does, however, make it possible to obtain an estimate of $\frac{\partial SSP_{it}}{\partial R_{it}}$.

In practice, $\frac{\partial SSP_{it}}{\partial R_{it}}$ is not expected to be much larger than zero for either sector. Unless there is a substantial break in policy the marginal effect of an additional unit of unconditional resources on education and health expenditure should be close to the average share of unconditional resources spent in the education and health sectors. As a very rough guide, if I proxy this share by the share of public education and health expenditure in total government expenditure, then for total government expenditure in the range of 20 to 30% of GDP the figures in table 1 suggest a marginal effect of unconditional resources of around 0.13-0.2 for education expenditure and 0.07-0.1 for health expenditure. Devarajan, Rajkumar, and Swaroop (2007) estimate the effect of unconditional resources on public education spending to be 0.12, whereas for public health expenditure the marginal effect is 0.04. Feyzioglu, Swaroop, and Zhu (1998) find even smaller effects of 0.08 for education expenditure and 0.02 for health expenditure. The assumption that $\frac{\partial SSP_{it}}{\partial R_{it}} = 0$ is therefore unlikely to have a significant impact on the conclusions drawn from the estimated coefficients and the full fungibility tests. Nonetheless, the reader is advised to keep in mind that this assumption should lead to a slight increase in the probability of rejecting the null hypothesis of full fungibility.

4 Results

4.1 Static panel results

Table 3 shows results from the OLS and fixed effects (FE) estimation of equation (20), with total donor-reported education or health aid as the main regressor of interest. The hypotheses tests for no and full fungibility in this table are therefore carried out under the assumption that education and health aid are completely on-budget.

All reported standard errors are robust to heteroscedasticity and are clustered at the country level, thereby allowing for serial correlation in the error term (Arellano, 1987). If not taken into account, serial correlation can lead to a serious underestimation of standard errors and lead one to over-reject the null hypothesis under consideration (Bertrand, Duflo, and Mullainathan, 2004; Kézdi, 2004). In addition, Stock and Watson (2008) show that the usually applied heteroscedasticity-robust estimator of the variance-covariance matrix (see White, 1980) is inconsistent in FE estimation with fixed T, whereas the cluster-robust estimator is still consistent. Bertrand, Duflo, and Mullainathan (2004) and Kézdi (2004) argue the cluster-robust estimator works well as long as the number of clusters does not become very small.

In both OLS and FE estimation education aid has no discernible correlation with public education expenditure and the null of no fungibility is strongly rejected. Health aid, by contrast, is positively correlated with public health expenditure in both OLS and FE and this effect is estimated precisely enough to reject the null of full fungibility as well as the null of no fungibility. However, the size of the FE coefficient of health aid is small: an increase in health aid of 1% of GDP is associated with an increase in public health expenditure of

Table 3: Total education and health aid

	Public education exp.		Public health exp.	
	OLS	FE	OLS	FE
Education aid	0.047	0.0042		
	(0.082)	(0.068)		
Health aid			0.47***	0.26**
			(0.18)	(0.12)
General aid	-0.0032	0.032	0.016	0.0037
	(0.053)	(0.029)	(0.030)	(0.019)
Support to NGOs	-0.41	-0.38*	-0.13	-0.18**
	(0.33)	(0.21)	(0.17)	(0.091)
Other non-education aid	0.0026	-0.0041		
	(0.022)	(0.018)		
Other non-health aid			0.0084	-0.012
			(0.017)	(0.012)
GDP per capita	0.085	0.26*	0.17***	0.14*
	(0.059)	(0.14)	(0.048)	(0.085)
GDP per capita growth	-0.049***	-0.028***	-0.025**	-0.020***
	(0.016)	(0.0093)	(0.012)	(0.0074)
Urbanisation	-0.010	0.080	0.0026	0.056*
	(0.0083)	(0.056)	(0.0053)	(0.033)
Trade	0.015***	-0.014**	0.010***	-0.0075*
	(0.0038)	(0.0068)	(0.0031)	(0.0041)
PV debt	-0.0038	-0.0025	0.00025	0.000032
	(0.0035)	(0.0017)	(0.0022)	(0.00056)
Public debt service	0.050	-0.063***	-0.040**	-0.024**
	(0.062)	(0.022)	(0.019)	(0.012)
R^2	0.178	0.207	0.294	0.171
Hausman		0.000		0.000
$\beta \leqslant 0$	0.285	0.475	0.005	0.019
$\beta \geqslant 1$	0.000	0.000	0.002	0.000
Countries	108	108	108	108
Observations	1082	1082	1087	1087

Note: OLS and fixed effects (FE) results, annual data, 1990-2003. All regressions include time dummies, coefficients not reported. Heteroscedasticity-robust standard errors, clustered by country, in brackets. $*$, $**$, and $***$ denote significance at 10, 5 and 1%, respectively. In the case of FE estimation, R^2 refers to the within R^2. Hausman shows the p-value of a generalised Hausman test of the null hypothesis that η_i is uncorrelated with the regressors. $\beta \leqslant 0$ ($\beta \geqslant 1$) is the p-value for the test of full (no) fungibility for total sectoral aid.

only 0.26% of GDP. So, on the basis of this, one would still conclude that health aid is mostly fungible.

As explained in section 2.2, the coefficients reported in table 3 are likely to overestimate the extent of fungibility because the presence of off-budget aid in the donor-based aid measure pulls down the estimated marginal effect of sectoral aid on public sectoral expenditure. Table 4 shows results from the estimation of equation (21), where sectoral aid is further partitioned into four prefix codes. This allows a separate evaluation of the effects of on-and off-budget sectoral aid and an implementation of the more precise fungibility tests described in table 2, using SP aid as a measure of on-budget aid and TC as a proxy for off-budget aid.

The further disaggregation of sectoral aid markedly changes results. In both sectors the marginal effect of SP aid in the FE model is close to 1, suggesting the bulk of sector programme aid is used in the intended sector. Full fungibility can be rejected whereas the null of no fungibility cannot be rejected. The effect of TC

Table 4: Disaggregated education and health aid

	Public education exp.		Public health exp.	
	OLS	FE	OLS	FE
Education IP	0.091	0.12		
	(0.25)	(0.12)		
Education SP	2.53*	1.21**		
	(1.35)	(0.55)		
Education TC	0.032	-0.0070		
	(0.10)	(0.082)		
Education ONM	0.14	0.021		
	(0.21)	(0.19)		
Health IP			0.40	0.20
			(0.34)	(0.21)
Health SP			1.19*	0.84***
			(0.60)	(0.31)
Health TC			-0.12	0.0067
			(0.35)	(0.32)
Health ONM			0.74**	0.41*
			(0.36)	(0.23)
General aid	-0.0012	0.031	0.023	0.0055
	(0.051)	(0.029)	(0.031)	(0.019)
Support to NGOs	-0.56*	-0.39**	-0.15	-0.16
	(0.30)	(0.19)	(0.16)	(0.11)
Other non-education aid	-0.0081	-0.0055		
	(0.022)	(0.018)		
Other non-health aid			0.014	-0.013
			(0.017)	(0.011)
GDP per capita	0.084	0.29*	0.17***	0.15*
	(0.060)	(0.15)	(0.048)	(0.085)
GDP per capita growth	-0.051***	-0.029***	-0.028**	-0.021***
	(0.015)	(0.0091)	(0.011)	(0.0072)
Urbanisation	-0.0089	0.085	0.0026	0.055*
	(0.0081)	(0.055)	(0.0053)	(0.031)
Trade	0.016***	-0.013**	0.011***	-0.0071*
	(0.0039)	(0.0067)	(0.0032)	(0.0040)
PV debt	-0.0040	-0.0027*	-0.000074	-0.000092
	(0.0034)	(0.0016)	(0.0021)	(0.00059)
Public debt service	0.052	-0.065***	-0.039**	-0.022*
	(0.062)	(0.021)	(0.019)	(0.011)
R^2	0.187	0.215	0.302	0.183
Hausman		0.000		0.000
$\beta_{SP} \leqslant 0$	0.032	0.015	0.026	0.004
$\beta_{SP} \geqslant 1$	0.870	0.645	0.621	0.307
$\beta_{TC} \leqslant -1$	0.000	0.000	0.006	0.001
$\beta_{TC} \geqslant 0$	0.621	0.466	0.363	0.508
Countries	108	108	108	108
Observations	1082	1082	1087	1087

Note: OLS and fixed effects (FE) results, annual data, 1990-2003. All regressions include time dummies, coefficients not reported. Heteroscedasticity-robust standard errors, clustered by country, in brackets. $*$, $**$, and $***$ denote significance at 10, 5 and 1%, respectively. In the case of FE estimation, R^2 refers to the within R^2. Hausman shows the p-value of a generalised Hausman test of the null hypothesis that η_i is uncorrelated with the regressors. $\beta_{SP} \leqslant 0$ ($\beta_{SP} \geqslant 1$) and $\beta_{TC} \leqslant -1$ ($\beta_{TC} \geqslant 0$) are p-values for the test of full (no) fungibility for sector programme aid and technical cooperation, respectively.

is close to zero in both sectors and the null of full fungibility is strongly rejected. The no fungibility hypothesis cannot be rejected, indicating there is no evidence that sectoral TC displaces a recipient government's own expenditure in either sector. The TC effect is similar in OLS, while the coefficients of SP aid become larger but are also estimated less precisely. The larger SP coefficients in OLS may indicate that time-invariant unobserved variables are positively correlated with both SP aid and sectoral public expenditure. In FE estimation the coefficients are identified from the within-country variation in the data, which reduces the problem of omitted variables in the case where such variables remain relatively fixed over time.

For the FE models in tables 3 and 4, a generalised Hausman test that allows for heteroscedasticity and serial correlation is reported (Arellano, 1993; Wooldridge, 2002, pp. 290-291).[19] The null hypothesis that η_i is uncorrelated with the regressors is always rejected, suggesting FE should be preferred over RE.

Overall, the results in table 4 illustrate that a failure to deal properly with the presence of off-budget sectoral aid may yield misleading conclusions with respect to fungibility. Once on- and off-budget aid are separated and the fungibility of each is assessed against its appropriate benchmark, the FE results suggest there is little if any fungibility. This conclusion is robust to a large number of changes to the specification. I replace the PV of debt by a non-PV measure of long-term external public and publicly guaranteed debt expressed as a percentage of GDP (from World Bank, 2006b). I also add, each in turn, two different measures of the PV of debt relief constructed by Depetris Chauvin and Kraay (2005) to the model.[20] As debt relief is often linked to higher education and health expenditure one might expect it has a larger positive impact on public education and health expenditure than the effect achieved by a reduction in debt or debt service that comes about through other means than debt relief. If this is the case, we would expect to see a positive effect of debt relief even after controlling for the level of debt and debt service. I do not find evidence for this. Even without controlling for debt or debt service I find no effect of the PV of debt relief on public education or health expenditure. I further include GDP per capita in log form rather than in thousands of dollars or add (one at a time) control variables for female labour force participation, the age dependency ratio and the birth rate (all from World Bank, 2006c), measures of corruption, the rule of law and bureaucratic quality from the International Country Risk Guide (The Political Risk Services Group, 2008), and measures of democracy taken from Polity IV (Marshall and Jaggers, 2007). In all cases the results are qualitatively unchanged. The only exception is that, when the ICRG measures are added, the coefficient of health TC drops to around -0.25 and I can reject the null hypothesis of no fungibility, implying partial (but still quite low) fungibility of health TC.

[19]This test is carried out in Stata using the xtoverid command (Schaffer and Stillman, 2006).

[20]I am grateful to Nicolas Depetris Chauvin for sharing this data.

4.1.1 Influential observations

Especially given the limited variation in education and health SP aid and, to a lesser extent, in education and health TC, one worry might be that the effects of these variables are driven by a small number of observations, which would create problems when generalising results. In the context of cross-country growth regressions, Temple (2000) argues more care should be taken to gauge to what extent results are driven by outliers. This critique is equally relevant for fungibility studies. In Devarajan, Rajkumar, and Swaroop (2007), for instance, the removal of Botswana turns a significant negative effect of education loans on public education expenditure into a strong positive effect. While, in the models of this paper, the addition of extra control variables generally leaves conclusions unchanged, in a few instances, especially when the addition of an extra variable leads to a large drop in sample size, the point estimates on the variables of interest shift by a relatively large amount. When this occurs it is always due to the change in the sample composition and not because the additional control variable takes away some of the explanatory power of sectoral SP aid or TC.[21]

As a first attempt to judge the sensitivity of results to outliers, I re-estimate equation (21) in log-linear form. Taking the natural logarithm of all variables compresses the upper tail and is therefore likely to reduce the influence of observations with larger values of education and health SP aid or TC on the estimated regression line. For the public expenditure, aid and debt variables, I add 1 before taking the log in order to deal with zero values. Since GDP per capita growth can be negative I enter it without taking its log.[22] The estimated coefficients now represent elasticities and the marginal effect of the variables of interest (here illustrated for SP aid), evaluated at the sample means, is calculated as follows:

$$\widehat{\beta}_{SP} = \frac{\widehat{\partial SSP}}{\partial SAIDSP} = \widehat{\zeta}_{\beta_{SP}}\left(\frac{1+\overline{SSP}}{1+\overline{SAIDSP}}\right) \tag{22}$$

where $\widehat{\zeta}_{\beta_{SP}}$ is the estimated elasticity and $\overline{SSP}$ and $\overline{SAIDSP}$ are sample averages of public sectoral expenditure and sectoral SP aid. Table 5 displays marginal effects for SP aid and TC calculated in this manner (full estimation results of the log-linear model are available on request). Results are very similar to the ones obtained in the linear model. In both sectors the effect of TC is close to zero and SP aid has a marginal effect on public expenditure that is close to 1. Full fungibility is rejected across the board, while the null of no fungibility cannot be rejected in any of the cases.

As a more direct and arguably superior approach to check for the impact of influential observations, I re-estimate equation (21) dropping one country at a time. Figure 3 shows the resulting distribution of the estimated SP aid and TC coefficients. The marginal effect of TC is more stable than that of SP aid in both sectors, which

[21] The most extreme deviation occurs when the birth rate is added, in which case the sample size in the health model falls to 612 and the effect of health SP aid rises to 1.34.

[22] GDP per capita is expressed in constant 2000 international dollars instead of *thousands* of constant 2000 international dollars before taking its log.

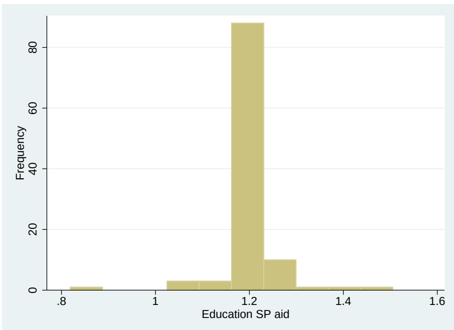
Table 5: Disaggregated education and health aid, marginal effects of the log-linear model

	Public education exp.	Public health exp.
$\widehat{\beta}_{SP}$	1.342	1.092
$\beta_{SP} \leqslant 0$	0.005	0.006
$\beta_{SP} \geqslant 1$	0.750	0.585
$\widehat{\beta}_{TC}$	0.0522	0.0602
$\beta_{TC} \leqslant -1$	0.000	0.000
$\beta_{TC} \geqslant 0$	0.632	0.591

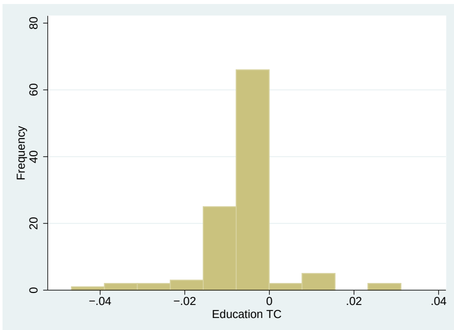
Note: marginal effects, calculated at the sample means, based on the fixed effects estimation of equation (21) in log-linear form (see main text for details). Annual data, 1990-2003. All regressions include time dummies and the standard set of control variables (coefficients not reported) and are estimated with heteroscedasticity-robust standard errors, clustered by country. $\beta_{SP} \leqslant 0$ ($\beta_{SP} \geqslant 1$) and $\beta_{TC} \leqslant -1$ ($\beta_{TC} \geqslant 0$) are p-values for the test of full (no) fungibility for sector programme aid and technical cooperation, respectively.

is consistent with the more limited extent of variation in the SP aid variables. A small number of countries induce fairly big changes in the effect of SP aid. When Lesotho is dropped, for instance, the effect of education SP aid falls to 0.82. When Tonga is excluded it increases to 1.51. In addition, a few other countries cause more moderate shifts in the SP aid coefficients when they are removed from the sample. In contrast, the distribution of the estimated coefficient of education TC has a much smaller range. For health TC two countries have a sizeable impact on the estimated coefficient when they are left out of the sample, but the remainder of the distribution is again much narrower.[23]

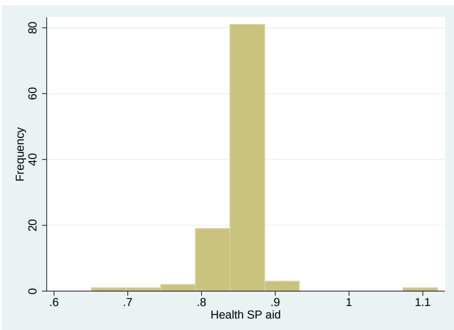
Figure 3: Distribution of coefficients when dropping one country at a time



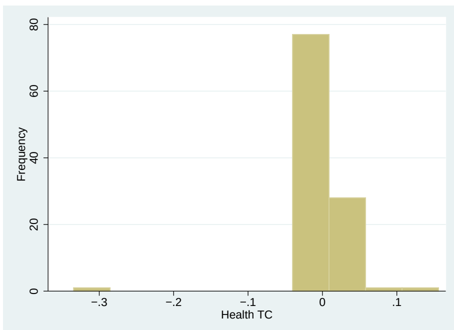
(a) Education SP aid



(b) Education TC



(c) Health SP aid



(d) Health TC

To examine how sensitive results are to the removal of countries that appear to exert an undue influence on

[23] Without Eritrea, the estimated effect of education TC becomes -0.33. Without Guinea-Bissau, the effect is 0.16. These two countries are also the ones that have the biggest impact on the estimated health SP aid coefficient.

Table 6: Disaggregated education and health aid, reduced sample

	Public education exp.		Public health exp.	
	FE	FD	FE	FD
Education IP	0.22	0.34^{***}		
	(0.15)	(0.12)		
Education SP	0.83^{**}	-0.34		
	(0.34)	(0.55)		
Education TC	0.024	-0.070		
	(0.059)	(0.046)		
Education ONM	-0.25	-0.044		
	(0.24)	(0.11)		
Health IP			0.17	-0.19^{*}
			(0.19)	(0.11)
Health SP			0.83^{**}	-0.19
			(0.36)	(0.24)
Health TC			-0.15	-0.040
			(0.20)	(0.10)
Health ONM			0.31^{*}	0.095
			(0.17)	(0.12)
General aid	0.027	0.00092	0.0082	-0.0074
	(0.020)	(0.015)	(0.018)	(0.011)
Support to NGOs	-0.48	-0.17	-0.055	-0.025
	(0.31)	(0.23)	(0.14)	(0.073)
Other non-education aid	0.00046	0.0049		
	(0.016)	(0.010)		
Other non-health aid			-0.019^{*}	0.0039
			(0.011)	(0.0047)
GDP per capita	0.22^{*}	-0.058	0.093	0.13^{**}
	(0.12)	(0.088)	(0.071)	(0.063)
GDP per capita growth	-0.019^{***}	-0.0079^{**}	-0.015^{***}	-0.011^{***}
	(0.0045)	(0.0031)	(0.0043)	(0.0031)
Urbanisation	0.039	0.0033	0.019	0.017
	(0.045)	(0.064)	(0.025)	(0.026)
Trade	-0.0035	-0.0025	-0.0013	0.00046
	(0.0041)	(0.0037)	(0.0023)	(0.0021)
PV debt	-0.0055^{**}	-0.0038	-0.00026	-0.00017
	(0.0025)	(0.0038)	(0.00056)	(0.00072)
Public debt service	-0.059^{***}	-0.024^{*}	-0.019	-0.0031
	(0.019)	(0.012)	(0.012)	(0.0057)
R^2	0.183	0.062	0.135	0.051
Hausman	0.000		0.000	
$\beta_{SP} \leqslant 0$	0.008	0.731	0.012	0.781
$\beta_{SP} \geqslant 1$	0.307	0.008	0.313	0.000
$\beta_{TC} \leqslant -1$	0.000	0.000	0.000	0.000
$\beta_{TC} \geqslant 0$	0.658	0.066	0.239	0.347
Countries	94	94	103	102
Observations	921	819	1024	912

Note: fixed effects (FE) and first-differenced OLS (FD) results, annual data, 1990-2003, reduced sample. All regressions include time dummies, coefficients not reported. Heteroscedasticity-robust standard errors, clustered by country, in brackets. $*$, $**$, and $***$ denote significance at 10, 5 and 1%, respectively. In the case of FE estimation, R^2 refers to the within R^2. Hausman shows the p-value of a generalised Hausman test of the null hypothesis that η_i is uncorrelated with the regressors. $\beta_{SP} \leqslant 0$ ($\beta_{SP} \geqslant 1$) and $\beta_{TC} \leqslant -1$ ($\beta_{TC} \geqslant 0$) are p-values for the test of full (no) fungibility for sector programme aid and technical cooperation, respectively.

the coefficients of interest, in both sectors I calculate a $DFBETA_i$ influence statistic for SP aid and TC for each country i (I again illustrate using SP aid as an example):

$$DFBETA^i_{SP} = \frac{\widehat{\beta}^i_{SP} - \widehat{\beta}_{SP}}{\widehat{SE}_{\widehat{\beta}^i_{SP}}} \tag{23}$$

where $\widehat{\beta}_{SP}$ is the estimated coefficient in the full sample, $\widehat{\beta}^i_{SP}$ is the estimate when country i is dropped and $\widehat{SE}_{\widehat{\beta}^i_{SP}}$ is the estimated standard error of the coefficient in the model without country i (see e.g. Bollen and Jackman, 1990). Belsley, Kuh, and Welsch (1980) suggest a size-adjusted cut-off of $2/\sqrt{N}$ (N in this case being the number of countries) to identify observations that require special attention. Hence, in each sector, I drop countries for which $|DFBETA^i_{SP}|$ and/or $|DFBETA^i_{TC}|$ exceeds $2/\sqrt{N}$ to investigate whether the results in table 4 are driven by only a few aid recipients. This procedure removes 14 countries in the education sector and 5 in the health sector. Table 6 shows results from estimating equation (21) on this reduced sample.

Generally speaking, FE results in the reduced sample are very similar to those in the full sample. The effect of TC in both sectors remains close to zero and full fungibility is easily rejected. The effect of education SP aid drops quite severely to 0.83, which is also the size of the almost unchanged coefficient of health SP aid, but in both cases full fungibility is still rejected. The conclusions from table 4, namely that the fungibility of education and health SP aid and TC is very limited, therefore continue to hold after excluding those countries that exert the largest influence on the estimated coefficients of interest. This suggests the patterns uncovered are not solely driven by the particular experience of a small number of aid recipients.

While FE is often the estimator of choice in cross-country empirical studies, to interpret the coefficients in a causal way requires a potentially strong assumption of strict exogeneity. Strict exogeneity entails the right hand side variables are uncorrelated with the idiosyncratic error in *any* time period. Focusing on the variables of interest, this implies assuming $\mathrm{E}(SAIDSP_{is}\epsilon_{it}) = \mathrm{E}(SAIDTC_{is}\epsilon_{it}) = 0$ for all $s, t = 1, \ldots, T$. This assumption would be violated if, for instance, the allocation of education (health) SP aid and TC is partly determined on the basis of past or current values of public education (health) expenditure.

In fact, table 6 contains some evidence indicating strict exogeneity is unlikely to hold. If a first-differenced version of equation (21) is estimated with OLS (second and fourth column, labelled FD) the effect of SP aid is markedly different from its FE estimate. In both sectors the estimated SP aid coefficient becomes negative. Full fungibility can now no longer be rejected, whereas the null of no fungibility can be rejected. This stark difference between FE and FD estimates of the SP aid coefficients suggests a violation of the strict exogeneity assumption, as such a violation causes both FE and FD to be inconsistent as well as generally to have different probability limits (Wooldridge, 2002, p. 284-285; also see Laporte and Windmeijer, 2005). At the same time, the effect of TC is very much the same in the first-differenced model. There is some evidence of a negative effect of TC, especially in the education sector where the no fungibility hypothesis can be rejected at a 10% significance level, but any displacement of sectoral public expenditure is minimal. Hence, the conclusion that the fungibility of TC is limited is confirmed in the FD model.

A second indication that the FE model is potentially misspecified comes from a serial correlation test on the idiosyncratic errors. Under the null of no serial correlation, residuals in the first-differenced model should have an autocorrelation of -0.5, so a Wald test of this hypothesis can be performed to test for the presence of serial correlation in ϵ_{it} (Wooldridge, 2002, p. 283; Drukker, 2003).[24] For both sectors I reject the null of no serial correlation at a less than 1% significance level. While clustering standard errors on the recipient country should ensure that inference is valid, the presence of serial correlation in ϵ_{it} may indicate the model is dynamically misspecified, which would again render the FE estimates inconsistent.

4.2 Dynamic panel results

In the remainder of this paper I relax the strict exogeneity assumption by employing a system GMM estimator.[25] This estimator also allows some dynamics in the determination of public education and health expenditure. In particular, it enables the consistent estimation of a more general model that includes a lagged dependent variable:

$$\begin{aligned} SSP_{it} = {} & \alpha SSP_{i,t-1} + \beta_{IP} SAIDIP_{it} + \beta_{SP} SAIDSP_{it} + \beta_{TC} SAIDTC_{it} \\ & + \beta_{ONM} SAIDONM_{it} + \boldsymbol{\gamma} \boldsymbol{A}_{it} + \boldsymbol{\delta} \boldsymbol{X}_{it} + \lambda_t + \eta_i + \epsilon_{it} \end{aligned} \tag{24}$$

Public education and health expenditure are likely to be persistent and modelling this persistence could be important to recover a consistent estimate of the effect of sectoral aid on public sectoral spending. As I show below, the inclusion of a lagged dependent variable removes the serial correlation in ϵ_{it}.

Building on an idea in Anderson and Hsiao (1981, 1982), the frequently employed difference GMM estimator (Holtz-Eakin, Newey, and Rosen, 1988; Arellano and Bond, 1991) differences equation (24) to remove the fixed effect and uses suitably lagged levels (or differences) of the dependent variable and the right hand side variables as instruments for the differenced equation. Under the assumption that SSP_{i1} is predetermined, meaning that it is uncorrelated with subsequent errors $\epsilon_{i2}, \ldots, \epsilon_{iT}$, and that there is no serial correlation in ϵ_{it}, $SSP_{i,t-2}$ and further lags $SSP_{i,t-3}, \ldots, SSP_{i1}$ are uncorrelated with $\Delta\epsilon_{it}$, making them valid instruments for the first-differenced equation at time $t = 3, \ldots, T$. A similar logic can be applied to the other right hand side variables. For instance, if a variable X_{it} is endogenous, in the sense that it is correlated with current and past disturbances but not with future disturbances ($\mathrm{E}(X_{it}\epsilon_{is}) = 0$ for $t < s$), lags $X_{i,t-l}$ for $l \geqslant 2$ are valid instru-

[24] I carry out this test in Stata using the xtserial command.

[25] For excellent discussions of difference and system GMM estimators, see Bond (2002) and Roodman (2009a).

ments for the differenced equation. Under the stricter assumption that X_{it} is predetermined ($\mathrm{E}(X_{it}\epsilon_{is}) = 0$ for $t \leqslant s$) lags $X_{i,t-l}$ for $l \geqslant 1$ are valid instruments.[26]

This leads to a set of population moment conditions that can be exploited within a Generalised Method of Moments (GMM) framework (Hansen, 1982). The GMM estimator minimises a weighted quadratic distance of the sample analogues of the population moment conditions. In an asymptotically efficient two-step GMM estimator, the moment conditions are weighted by a consistent estimate of their covariance matrix, which, in turn, is based on estimates of the first-differenced errors from an initial consistent estimator. While the two-step GMM estimator increases efficiency, its standard errors also suffer a severe downward finite sample bias (Arellano and Bond, 1991; Windmeijer, 2005). Windmeijer (2005) makes an approximate correction available that deals with this bias and that is shown to work well in Monte Carlo simulations. Hence, all results I present below are from a two-step GMM estimator that employs Windmeijer's (2005) correction, which is available in Stata's xtabond2 command (Roodman, 2009a).

A potential problem with the difference GMM estimator is that lagged levels might provide weak instruments for current first differences, especially if variables are persistent over time. This weak instrument problem leads to imprecision and, more importantly, a serious finite sample bias in the difference GMM estimator (Arellano and Bover, 1995; Blundell and Bond, 1998). Arellano and Bover (1995) and Blundell and Bond (1998) therefore suggest also employing the equation in levels, using suitably lagged differences as instruments, in what has become known as the system GMM estimator. For instance, for an endogenous variable X_{it}, an additional $T-2$ non-redundant moment conditions $\mathrm{E}(\Delta X_{i,t-1}(\eta_i + \epsilon_{it})) = 0$ for $t = 3, \ldots, T$ are available. For this approach to be valid, it is necessary that first-differences of the variables are uncorrelated with η_i. This, in turn, can be traced back to an assumption on the initial conditions (Blundell and Bond, 1998; Roodman, 2009b). For SSP_{it}, for instance, this entails that, conditional on the other regressors, the initial deviation from its long-run mean, $SSP_{i1} - \eta_i/(1-\alpha)$, is uncorrelated with the fixed effect η_i and therefore uncorrelated with the level of the long-run mean $\eta_i/(1-\alpha)$. This holds automatically if the same process has been generating the data for a long enough period before the sample starts.

An issue in the application of the system GMM estimator, already noted by Sargan (1958) in the context of instrumental variables estimation, is the risk of overfitting (Roodman, 2009b). As GMM creates an instrument for each variable in each time period for each lag distance, a large number of instruments are generated, especially if T is relatively long. As a result, the instrumentation strategy might fail to isolate the exogenous component of the right-hand side variables. Roodman (2009b) illustrates with a useful 2SLS analogy: if the number of instruments equals the number of observations, the first stage yields an R^2 of one and 2SLS gives the same result as OLS because it fails to extract the exogenous component of the endogenous variable. Hence,

[26] If X_{it} is strictly exogenous the full time series X_{it} for $t = 1, \ldots, T$ can be used as instruments.

overfitting biases coefficients towards those in the non-instrumented equation. In addition, the Hansen J-test of over-identifying restrictions loses power and therefore tends to provide a false sense of comfort in the validity of the instruments. I reduce the risk of overfitting by limiting the number of instruments used. That is, I restrict the maximum number of lags of the level variables used as instruments for the differenced equation. In addition, I collapse the instrument set, by creating an instrument for each variable and lag distance rather than for each variable, time period, and lag distance (Roodman, 2009a,b).

Two-step system GMM results using Windmeijer's (2005) correction for the standard errors are shown in tables 7 and 8 for the education and health sector, respectively. Since education (health) aid may be purposefully allocated towards countries with increasing or decreasing contemporaneous public education (health) expenditure, all education (health) aid prefix code variables are treated as endogenous. One might expect donors to give more sectoral SP aid to countries that are developing better sectoral policies and these would typically be the countries with increasing public sectoral expenditure. TC, by contrast, is intended mainly to help countries craft such policies and design effective service delivery mechanisms. As such, donors might allocate TC predominantly to countries with decreasing sectoral expenditure.

Support to NGOs is also allowed to be endogenous. Donors may, for instance, decide to channel more aid through NGOs if the recipient government fails to provide education and health services to all or part of the population. Increased trade might raise the demand for public expenditure. On the other hand, however, trade openness may act as a constraint on the expansion of government if the latter reduces trade competitiveness (e.g. via higher payroll taxes or export taxes). As this competitiveness loss hurts more in an open economy, higher external trade may make governments more reluctant to raise public expenditure. This reasoning implies, however, that there is potential feedback from current and past public expenditure to trade, so trade as well is treated as endogenous. All other variables are treated as predetermined. Time dummies are treated as strictly exogenous and therefore added to the instrument matrix without any transformation.[27]

Results are presented for a number of instrument configurations. The first column in both tables uses only a single lag of each variable to instrument the differenced equation, until the fourth column which uses four lags of each variable. This allows to examine how results change when the number of instruments is reduced, which limits the risk that overfitting is driving the results. As discussed above, the instrument set is collapsed to further reduce the number of instruments.

The short-run effect of SP aid in both sectors hovers around zero but is volatile across the different instrument configurations and is estimated imprecisely. As a result, neither the null of full fungibility nor the null of no fungibility can typically be rejected at conventional significance levels. This volatility and imprecision

[27]Time dummies are used as instruments for the levels equation only, as their additional use as instruments for the differenced equation is superfluous.

Table 7: Disaggregated education aid, reduced sample, system GMM

	Public education exp.			
	1 lag	2 lags	3 lags	4 lags
Lag Public education exp.	0.76***	0.75***	0.75***	0.73***
	(0.10)	(0.072)	(0.069)	(0.073)
Education IP	0.26	0.43	0.078	0.041
	(0.44)	(0.51)	(0.32)	(0.51)
Education SP	-0.55	-0.23	0.057	-0.066
	(1.24)	(1.02)	(0.96)	(0.70)
Education TC	0.026	-0.027	-0.075	-0.081
	(0.11)	(0.12)	(0.080)	(0.095)
Education ONM	-0.18	-0.23	-0.24	-0.33
	(0.41)	(0.40)	(0.29)	(0.28)
General aid	-0.035	-0.037	-0.047**	-0.036*
	(0.021)	(0.026)	(0.022)	(0.020)
Support to NGOs	1.64	1.57**	1.67**	1.49**
	(1.07)	(0.78)	(0.65)	(0.70)
Other non-education aid	-0.00021	-0.0096	-0.0027	0.0095
	(0.038)	(0.034)	(0.027)	(0.026)
GDP per capita	0.0013	-0.014	-0.037	-0.055
	(0.047)	(0.060)	(0.047)	(0.048)
GDP per capita growth	-0.0072	-0.013	-0.016**	-0.018**
	(0.0081)	(0.0082)	(0.0073)	(0.0090)
Urbanisation	0.0032	0.0018	0.0041	0.0090
	(0.0096)	(0.013)	(0.010)	(0.012)
Trade	0.0017	0.0036	0.0056	0.0065
	(0.0054)	(0.0049)	(0.0044)	(0.0043)
PV debt	-0.0012	-0.0050	-0.0063	-0.0055
	(0.0086)	(0.0093)	(0.0057)	(0.0050)
Public debt service	-0.013	-0.020	-0.011	-0.016
	(0.031)	(0.026)	(0.022)	(0.019)
# instruments	41	55	69	83
$m1$	0.006	0.009	0.009	0.010
$m2$	0.341	0.353	0.348	0.365
Hansen	0.709	0.664	0.528	0.352
Hansen level	0.709	0.413	0.130	0.211
Hansen lagdep diff	0.526	0.126	0.302	0.240
Hansen lagdep level	0.469	0.403	0.560	0.893
Hansen predeterm	0.709	0.796	0.397	0.156
$\beta_{SP} \leqslant 0$	0.672	0.588	0.477	0.538
$\beta_{SP} \geqslant 1$	0.106	0.116	0.165	0.067
$\beta_{TC} \leqslant -1$	0.000	0.000	0.000	0.000
$\beta_{TC} \geqslant 0$	0.591	0.411	0.176	0.200
β_{SP}^{LR}	-2.34	-0.91	0.22	-0.25
$\beta_{SP}^{LR} \leqslant 0$	0.674	0.588	0.477	0.537
$\beta_{SP}^{LR} \geqslant 1$	0.260	0.321	0.419	0.317
β_{TC}^{LR}	0.11	-0.11	-0.30	-0.30
$\beta_{TC}^{LR} \leqslant -1$	0.011	0.034	0.019	0.034
$\beta_{TC}^{LR} \geqslant 0$	0.591	0.412	0.189	0.217
Countries	94	94	94	94
Observations	829	829	829	829

Note: two-step system GMM estimator using Windmeijer's (2005) correction, annual data, 1990-2003, reduced sample. Column titles identify the number of lags of each variable used to instrument the differenced equation. The instrument matrix is collapsed. All regressions include time dummies, coefficients not reported. $*$, $**$, and $***$ denote significance at 10, 5 and 1%, respectively. $m1$ and $m2$ show p-values for Arellano and Bond's (1991) tests of first- and second-order serial correlation in the differenced residuals. Hansen is the p-value for Hansen's (1982) J-test of instrument validity. P-values for difference-in-Hansen tests of the following subsets of instruments are reported: all instruments in the level equation (Hansen level), the lagged dependent variable in the differenced (Hansen lagdep diff) and levels (Hansen lagdep level) equation, and the predetermined variables (Hansen predeterm). $\beta_{SP} \leqslant 0$ ($\beta_{SP} \geqslant 1$) and $\beta_{TC} \leqslant -1$ ($\beta_{TC} \geqslant 0$) are p-values for the test of full (no) fungibility for sector programme aid and technical cooperation, respectively. β_{SP}^{LR} and β_{TC}^{LR} present estimated long-run effects. No and full fungibility tests of the long-run effects are indicated with superscript LR. Observations refers to the number of observations in the levels equation.

Table 8: Disaggregated health aid, reduced sample, system GMM

	Public health exp.			
	1 lag	2 lags	3 lags	4 lags
Lag Public health exp.	0.86***	0.86***	0.85***	0.85***
	(0.081)	(0.077)	(0.060)	(0.050)
Health IP	-0.070	0.15	0.18	0.25
	(0.36)	(0.42)	(0.35)	(0.30)
Health SP	-0.40	0.11	0.13	0.095
	(0.99)	(1.06)	(0.90)	(0.81)
Health TC	0.35	0.43	0.38	0.35
	(0.51)	(0.29)	(0.24)	(0.23)
Health ONM	0.54	0.099	-0.048	-0.010
	(0.51)	(0.41)	(0.34)	(0.29)
General aid	-0.033	-0.024	-0.021	-0.017
	(0.021)	(0.022)	(0.019)	(0.017)
Support to NGOs	0.32	0.30	0.061	-0.063
	(0.33)	(0.35)	(0.30)	(0.22)
Other non-health aid	-0.0066	-0.016	-0.013	-0.012
	(0.017)	(0.013)	(0.011)	(0.0090)
GDP per capita	0.057	0.056*	0.053**	0.052**
	(0.047)	(0.033)	(0.026)	(0.022)
GDP per capita growth	-0.013**	-0.0065	-0.0064	-0.0077
	(0.0062)	(0.0056)	(0.0054)	(0.0050)
Urbanisation	-0.0063	-0.0056	-0.0063*	-0.0070*
	(0.0056)	(0.0043)	(0.0037)	(0.0036)
Trade	0.0049*	0.0011	0.0012	0.0018
	(0.0027)	(0.0017)	(0.0015)	(0.0015)
PV debt	0.00042	0.0014	0.0012	0.0010*
	(0.00070)	(0.00099)	(0.00079)	(0.00055)
Public debt service	0.0043	-0.0014	0.0038	0.0021
	(0.012)	(0.016)	(0.015)	(0.013)
# instruments	41	55	69	83
$m1$	0.000	0.000	0.000	0.000
$m2$	0.228	0.329	0.386	0.395
Hansen	0.344	0.209	0.625	0.803
Hansen level	0.344	0.305	0.572	0.280
Hansen lagdep diff	0.689	0.607	0.734	0.895
Hansen lagdep level	0.598	0.497	0.762	0.795
Hansen predeterm	0.344	0.172	0.563	0.679
$\beta_{SP} \leqslant 0$	0.658	0.457	0.443	0.454
$\beta_{SP} \geqslant 1$	0.079	0.202	0.166	0.135
$\beta_{TC} \leqslant -1$	0.004	0.000	0.000	0.000
$\beta_{TC} \geqslant 0$	0.753	0.927	0.946	0.933
β_{SP}^{LR}	-2.81	0.82	0.88	0.64
$\beta_{SP}^{LR} \leqslant 0$	0.651	0.456	0.442	0.453
$\beta_{SP}^{LR} \geqslant 1$	0.300	0.490	0.492	0.474
β_{TC}^{LR}	2.41	3.09	2.64	2.33
$\beta_{TC}^{LR} \leqslant -1$	0.181	0.018	0.011	0.012
$\beta_{TC}^{LR} \geqslant 0$	0.741	0.945	0.953	0.943
Countries	102	102	102	102
Observations	922	922	922	922

Note: two-step system GMM estimator using Windmeijer's (2005) correction, annual data, 1990-2003, reduced sample. Column titles identify the number of lags of each variable used to instrument the differenced equation. The instrument matrix is collapsed. All regressions include time dummies, coefficients not reported. $*$, $**$, and $***$ denote significance at 10, 5 and 1%, respectively. $m1$ and $m2$ show p-values for Arellano and Bond's (1991) tests of first- and second-order serial correlation in the differenced residuals. Hansen is the p-value for Hansen's (1982) J-test of instrument validity. P-values for difference-in-Hansen tests of the following subsets of instruments are reported: all instruments in the level equation (Hansen level), the lagged dependent variable in the differenced (Hansen lagdep diff) and levels (Hansen lagdep level) equation, and the predetermined variables (Hansen predeterm). $\beta_{SP} \leqslant 0$ ($\beta_{SP} \geqslant 1$) and $\beta_{TC} \leqslant -1$ ($\beta_{TC} \geqslant 0$) are p-values for the test of full (no) fungibility for sector programme aid and technical cooperation, respectively. β_{SP}^{LR} and β_{TC}^{LR} present estimated long-run effects. No and full fungibility tests of the long-run effects are indicated with superscript LR. Observations refers to the number of observations in the levels equation.

carry over to the estimate of the long-run effect of education SP aid, $\widehat{\beta}_{SP}^{LR} = \widehat{\beta}_{SP}/(1-\widehat{\alpha})$.[28] Most likely, this emanates from the lack of variation in SP aid.

The effect of education TC is very close to zero in all columns and the null of full fungibility is always strongly rejected. No fungibility cannot be rejected and the point estimate suggests at most only minor displacement of public education expenditure by education TC in the short-run. Given the persistence in public education expenditure the estimate of the long-run effect of education TC is somewhat more negative (with -0.3 as the lowest estimate) but even in the long-run full fungibility is rejected and no fungibility is not rejected. In the health sector as well, full fungibility of TC in the short run is rejected across the board. In fact, health TC is found to have a positive effect though the estimate is never significantly different from zero. The average estimated LR effect is around 2.6 and, in all cases except column 1, full fungibility in the long-run can be rejected.

In both sectors the coefficient of the lagged dependent variable lies between its FE and OLS estimate, which is what we would expect from a consistent estimator.[29] Arellano and Bond's (1991) $m2$-test never rejects the null hypothesis of no second-order serial correlation in $\Delta\epsilon_{it}$ at a 10% significance level, while the $m1$ test rejects the null of no first-order serial correlation in the differenced residuals. This suggests ϵ_{it} is not serially correlated. The Hansen J-test never raises concerns with regard to instrument validity and neither do the difference-in-Hansen tests that are carried out for the instruments used in the level equation (Hansen level), the lagged dependent variable used as an instrument in the differenced (Hansen lagdep diff) and level (Hansen lagdep level) equation, and the predetermined variables (Hansen predeterm).

Finally, while (24) is a dynamic equation in the sense that it explicitly models the persistence in public education and health expenditure, it may still be restrictive in only allowing for contemporaneous effects of sectoral aid. Dates of the fiscal year employed by recipient governments may not necessarily coincide fully with the calendar years to which the aid data refer. Moreover, even if fiscal and calendar years overlap perfectly, aid that arrives late in the year may only elicit a public expenditure response in the following year. If I replace sectoral TC by its lag the effect of lagged education TC is mildly positive (between 0.05 and 0.1 in the short run and 0.19 and 0.36 in the long run) and full fungibility can be rejected in both the short and long run.[30] The impact of health TC is slightly smaller when the lag is substituted in for the contemporaneous variable (0.15-0.27 in the short run, 1.14-1.87 in the long run) and full fungibility in the short run is again rejected across the board. As in table 8, full fungibility in the long-run is rejected except in the model where only a single lag of the variables in levels is used to instrument the differenced equation.

[28] Standard errors for the long-run effect are computed using the delta method.

[29] The pooled OLS estimate of α in equation (24) is upward biased, while the FE estimate suffers from a downward bias (Nickell, 1981; Bond, 2002). Here, the bounds suggested by the OLS and FE estimate of α are 0.5-0.93 and 0.57-0.92 for the education and health sector, respectively.

[30] As the limited variation in SP aid prevents us from obtaining a precise estimate of its effect on public sectoral expenditure and drawing firm conclusions with respect to its degree of fungibility, in what follows I focus mainly on TC.

Table 9: Disaggregated education aid with TC lag added, reduced sample, system GMM

	Public education exp.			
	1 lag	2 lags	3 lags	4 lags
Lag Public education exp.	0.71***	0.72***	0.72***	0.71***
	(0.10)	(0.083)	(0.073)	(0.076)
Education IP	-0.044	-0.0088	-0.047	-0.18
	(0.38)	(0.46)	(0.32)	(0.47)
Education SP	-0.16	-0.077	-0.081	-0.13
	(0.95)	(1.02)	(0.81)	(0.65)
Education TC	-0.14	-0.097	-0.21*	-0.15
	(0.19)	(0.14)	(0.11)	(0.10)
Education TC $t-1$	0.20	0.11	0.18*	0.14
	(0.15)	(0.12)	(0.095)	(0.090)
Education ONM	-0.23	-0.34	-0.073	-0.31
	(0.39)	(0.39)	(0.32)	(0.32)
General aid	-0.0058	-0.024	-0.023	-0.018
	(0.032)	(0.033)	(0.028)	(0.026)
Support to NGOs	0.62	0.95	1.03	1.10
	(1.31)	(0.98)	(0.79)	(0.75)
Other non-education aid	0.029	0.019	0.019	0.019
	(0.026)	(0.024)	(0.020)	(0.020)
GDP per capita	0.019	0.00058	-0.022	-0.059
	(0.047)	(0.054)	(0.052)	(0.050)
GDP per capita growth	-0.0083	-0.013*	-0.018**	-0.015**
	(0.0083)	(0.0071)	(0.0068)	(0.0076)
Urbanisation	0.0060	0.0078	0.0095	0.012
	(0.011)	(0.011)	(0.011)	(0.011)
Trade	-0.0014	0.0027	0.0046	0.0052
	(0.0059)	(0.0042)	(0.0038)	(0.0038)
PV debt	-0.0012	-0.0024	-0.0035	-0.0047
	(0.0079)	(0.0091)	(0.0058)	(0.0046)
Public debt service	-0.0097	-0.028	-0.017	-0.011
	(0.032)	(0.029)	(0.025)	(0.022)
# instruments	41	55	69	83
$m1$	0.006	0.009	0.010	0.010
$m2$	0.276	0.302	0.302	0.320
Hansen	0.553	0.638	0.659	0.294
Hansen level	.	0.395	0.232	0.071
Hansen lagdep diff	0.139	0.378	0.516	0.554
Hansen lagdep level	0.259	0.524	0.951	0.244
Hansen predeterm	.	0.751	0.605	0.139
$\beta_{SP} \leqslant 0$	0.566	0.530	0.540	0.583
$\beta_{SP} \geqslant 1$	0.113	0.146	0.092	0.041
$\beta_{TC} \leqslant -1$	0.000	0.000	0.000	0.000
$\beta_{TC} \geqslant 0$	0.240	0.240	0.033	0.072
β^{LR}_{SP}	-0.54	-0.27	-0.29	-0.47
$\beta^{LR}_{SP} \leqslant 0$	0.566	0.530	0.540	0.582
$\beta^{LR}_{SP} \geqslant 1$	0.319	0.363	0.327	0.257
β^{LR}_{TC}	0.21	0.049	-0.13	-0.062
$\beta^{LR}_{TC} \leqslant -1$	0.000	0.001	0.002	0.001
$\beta^{LR}_{TC} \geqslant 0$	0.763	0.563	0.324	0.414
Countries	94	94	94	94
Observations	825	825	825	825

Note: two-step system GMM estimator using Windmeijer's (2005) correction, annual data, 1990-2003, reduced sample. Column titles identify the number of lags of each variable used to instrument the differenced equation. The instrument matrix is collapsed. All regressions include time dummies, coefficients not reported. $*$, $**$, and $***$ denote significance at 10, 5 and 1%, respectively. $m1$ and $m2$ show p-values for Arellano and Bond's (1991) tests of first- and second-order serial correlation in the differenced residuals. Hansen is the p-value for Hansen's (1982) J-test of instrument validity. P-values for difference-in-Hansen tests of the following subsets of instruments are reported: all instruments in the level equation (Hansen level), the lagged dependent variable in the differenced (Hansen lagdep diff) and levels (Hansen lagdep level) equation, and the predetermined variables (Hansen predeterm). $\beta_{SP} \leqslant 0$ ($\beta_{SP} \geqslant 1$) and $\beta_{TC} \leqslant -1$ ($\beta_{TC} \geqslant 0$) are p-values for the test of full (no) fungibility for sector programme aid and technical cooperation, respectively. β^{LR}_{SP} and β^{LR}_{TC} present estimated long-run effects. No and full fungibility tests of the long-run effects are indicated with superscript LR. Observations refers to the number of observations in the levels equation.

Table 10: Disaggregated health aid with TC lag added, reduced sample, system GMM

	Public health exp.			
	1 lag	2 lags	3 lags	4 lags
Lag Public health exp.	0.88***	0.86***	0.85***	0.84***
	(0.077)	(0.076)	(0.055)	(0.048)
Health IP	-0.15	0.086	0.20	0.22
	(0.48)	(0.39)	(0.33)	(0.27)
Health SP	-0.43	-0.30	-0.049	-0.085
	(1.33)	(1.26)	(0.89)	(0.77)
Health TC	-0.12	0.11	0.21	0.18
	(0.92)	(0.52)	(0.34)	(0.30)
Health TC $t-1$	0.13	0.21	0.19	0.19
	(0.28)	(0.22)	(0.21)	(0.17)
Health ONM	0.14	-0.11	-0.10	-0.092
	(0.62)	(0.37)	(0.30)	(0.27)
General aid	-0.033	-0.019	-0.021	-0.018
	(0.031)	(0.025)	(0.020)	(0.017)
Support to NGOs	0.32	0.44	0.23	0.097
	(0.51)	(0.41)	(0.32)	(0.23)
Other non-health aid	0.0020	-0.0081	-0.0100	-0.0087
	(0.023)	(0.014)	(0.012)	(0.0093)
GDP per capita	0.040	0.057	0.064**	0.055**
	(0.057)	(0.042)	(0.027)	(0.023)
GDP per capita growth	-0.011	-0.0063	-0.0069	-0.0069
	(0.0066)	(0.0061)	(0.0055)	(0.0048)
Urbanisation	-0.0053	-0.0052	-0.0068*	-0.0063*
	(0.0065)	(0.0050)	(0.0039)	(0.0037)
Trade	0.0047	0.00029	0.00055	0.00097
	(0.0039)	(0.0020)	(0.0018)	(0.0017)
PV debt	0.00053	0.0012	0.0011	0.00090
	(0.0012)	(0.0011)	(0.00089)	(0.00063)
Public debt service	0.011	0.0027	0.0071	0.0065
	(0.012)	(0.014)	(0.015)	(0.014)
# instruments	41	55	69	83
$m1$	0.000	0.000	0.000	0.000
$m2$	0.441	0.477	0.524	0.581
Hansen	0.253	0.149	0.662	0.851
Hansen level	.	0.263	0.575	0.317
Hansen lagdep diff	0.248	0.175	0.283	0.845
Hansen lagdep level	0.398	0.369	0.572	0.962
Hansen predeterm	.	0.139	0.625	0.782
$\beta_{SP} \leqslant 0$	0.626	0.595	0.522	0.543
$\beta_{SP} \geqslant 1$	0.143	0.152	0.120	0.082
$\beta_{TC} \leqslant -1$	0.171	0.019	0.000	0.000
$\beta_{TC} \geqslant 0$	0.447	0.580	0.732	0.723
β_{SP}^{LR}	-3.55	-2.24	-0.32	-0.54
$\beta_{SP}^{LR} \leqslant 0$	0.617	0.588	0.522	0.543
$\beta_{SP}^{LR} \geqslant 1$	0.351	0.374	0.410	0.379
β_{TC}^{LR}	0.078	2.30	2.60	2.35
$\beta_{TC}^{LR} \leqslant -1$	0.438	0.096	0.008	0.004
$\beta_{TC}^{LR} \geqslant 0$	0.505	0.819	0.960	0.968
Countries	102	102	102	102
Observations	919	919	919	919

Note: two-step system GMM estimator using Windmeijer's (2005) correction, annual data, 1990-2003, reduced sample. Column titles identify the number of lags of each variable used to instrument the differenced equation. The instrument matrix is collapsed. All regressions include time dummies, coefficients not reported. $*$, $**$, and $***$ denote significance at 10, 5 and 1%, respectively. $m1$ and $m2$ show p-values for Arellano and Bond's (1991) tests of first- and second-order serial correlation in the differenced residuals. Hansen is the p-value for Hansen's (1982) J-test of instrument validity. P-values for difference-in-Hansen tests of the following subsets of instruments are reported: all instruments in the level equation (Hansen level), the lagged dependent variable in the differenced (Hansen lagdep diff) and levels (Hansen lagdep level) equation, and the predetermined variables (Hansen predeterm). $\beta_{SP} \leqslant 0$ ($\beta_{SP} \geqslant 1$) and $\beta_{TC} \leqslant -1$ ($\beta_{TC} \geqslant 0$) are p-values for the test of full (no) fungibility for sector programme aid and technical cooperation, respectively. β_{SP}^{LR} and β_{TC}^{LR} present estimated long-run effects. No and full fungibility tests of the long-run effects are indicated with superscript LR. Observations refers to the number of observations in the levels equation.

Tables 9 and 10 present results for models that simultaneously allow for a lagged and contemporaneous effect of TC. In the education sector the effect of current TC is negative, while lagged TC has a positive impact. The long-run effect of education TC is close to zero and there is clear evidence against full fungibility in the long run.[31] The previously strong positive impact of support to NGOs on public education expenditure weakens substantially and the coefficient for this variable is now no longer significantly different from zero. In the health sector I estimate a positive effect of both contemporaneous and lagged TC. The LR effect has a similar size as before and full fungibility can generally be rejected. Only in the model that uses a single lag of the variables in levels to instrument for the differenced equation (column 1) is the coefficient of current education TC negative, which results in a long-run effect that is very close to zero. In this case, full fungibility in the long run can no longer be rejected. Generally, the Hansen and difference-in-Hansen statistics do not indicate that the instruments are invalid. As one would expect when carrying out a large number of instrument validity tests, a few of the p-values drop below 0.2 or even – in one case – below 0.1, but no clear pattern emerges that would lead us to reject the validity of the instruments.

5 Conclusion

This paper presents new empirical evidence to shed light on the thorny issue of foreign aid fungibility. I construct data on earmarked education and health aid disbursements that also, to some extent, distinguishes between on- and off-budget components of aid. Sector programme aid measures on-budget aid, while technical cooperation is used as a proxy for off-budget aid. I develop a simple analytical framework to illustrate how a failure to adequately deal with the presence of off-budget aid may have biased all previous estimates of foreign aid fungibility. A first noteworthy finding is that technical cooperation takes up a big share of education and health aid. This highlights that the bias from not dealing with off-budget aid in an adequate manner may be large.

Overall, I find little evidence to suggest that aid is fully fungible. In both sectors, even in the long run, technical cooperation leads to at most only a small displacement of a recipient's own public spending. While the effect of technical cooperation is robust across a range of models, the effect of sector programme aid is more volatile. In a static panel data model, fixed effects results suggest an approximately one-for-one increase in public sectoral expenditure in response to sector programme aid. However, when using system GMM to estimate a dynamic model, the effect of sector programme aid in both sectors becomes very imprecise, so that in the end no firm conclusions can be drawn with respect to the fungibility of sector programme aid.

[31] The opposite signs but similar magnitudes of the coefficients of education TC and its lag, and their similar standard errors, may arise from high correlations between them, in which case both coefficient estimates may be identified on the basis of only a small amount of individual variation in each variable and caution should be applied in their interpretation (Spanos and McGuirk, 2002; Roodman, 2008). As already discussed, however, a very similar long-run impact is found when either education TC or its lag are entered separately in the regression.

It should be emphasised that the result of less than full fungibility for earmarked education and health aid pertains specifically to technical cooperation. Since the extent to which investment projects and other aid are on- or off-budget is more uncertain and the lack of variation in sector programme aid precludes the precise estimation of its effect on public sectoral expenditure, the empirical analysis in this paper is not able to ascertain if the degree of fungibility differs by aid modality. As technical cooperation is the dominant modality in both sectors, however, it plays a large role in determining the overall degree of fungibility of total earmarked education and health aid.

The lack of fungibility of technical cooperation may be a consequence of effective donor conditionality. If donors are able to monitor the recipient government's spending, they may be able to credibly enforce the condition that the government does not cut back its planned expenditure after receiving technical cooperation. An additional reason to explain the low degree of fungibility found, that applies specifically to technical cooperation and less to the other aid types, is the observation made by Gramlich (1977) that heterogeneity in government expenditure might contribute to reduced fungibility. To the extent that governments in developing countries spend few of their resources on the type of goods and services that are provided by technical cooperation, it becomes impossible to reduce this class of expenditure by much, as it quickly hits a lower bound of zero. If, in addition, the substitutability between different types of expenditure in the recipient government's utility function is limited, low fungibility for technical cooperation may ensue. Finally, a lack of information on the recipient government's part may also reduce the degree of fungibility.

Appendix A Construction of the sectoral aid data

A.1 Creditor Reporting System

As already discussed in the main text, the OECD's (2002) Creditor Reporting System (CRS) disaggregates development assistance along a number of dimensions, including the sector or purpose of aid and the aid type or prefix code. Unfortunately, because CRS disbursements reported by most donors are incomplete in at least some years they need to be supplemented with additional information. This appendix describes in detail a data construction method that further makes use of two OECD Development Assistance Committee (DAC) data tables to construct more complete disaggregated aid disbursements.[32]

Starting from the CRS database, I download annual gross disbursements in millions of US$ for the period 1990-2004 for the following sectors:[33] education (DAC5 sector code 110), health (120), commodity aid/general programme assistance (500), action relating to debt (600), donor administrative costs (910), support to NGOs (920) and other sectors (the sum of all remaining sector codes).[34] These data are obtained in a recipient-donor-year (RDY) format, i.e. for each year it shows the amount of foreign aid transferred from each donor to each recipient. Education and health disbursements are further partitioned into four aid types or prefix codes: investment projects (IP), sector programme (SP) aid, technical cooperation (TC), and other (no mark) (ONM).[35]

The prefix codes are useful because, to some extent, they allow to separate on- and off-budget aid flows. Definitions of the prefix codes can be found in OECD (2002, p. 22) (also see OECD, 2000a, pp. 47-48): "free-standing technical cooperation is defined as financing of activities whose primary purpose is to augment the level of knowledge, skills, technical know-how or productive aptitudes of the population of aid recipient countries. … It includes the cost of personnel, training and research, as well as associated equipment and administrative costs" and mainly comes in the form of "supply of human resources (teachers, volunteers and experts) or action targeted on human resources (education, training, advice)" (OECD, 2000a, p. 47).[36] Sector programme aid "comprises contributions to carry out wide-ranging development plans in a defined sector such

[32] All data used in this appendix can be accessed via the OECD's International Development Statistics (IDS) online databases on aid and other resource flows at www.oecd.org/dac/stats/idsonline.

[33] In CRS, the sector is recorded using a 5-digit purpose code, the first 3 digits of which refer to the corresponding sector in DAC table 5 (see OECD, 2002, Annex 5, pp. 87-106). It is these 3 digits I focus on to demarcate sectors. DAC5 contains a disaggregation of total official development assistance along the same sectors and aid types as CRS, but in a donor-year format, not by recipient (see below for more information).

[34] Other sector aid consists of: population programmes (130), water supply and sanitation (140), government and civil society (150), other social infrastructure and services (160), economic infrastructure and services (200), production sectors (300), multisector/crosscutting (400), emergency assistance (700) and unallocated/unspecified (998).

[35] In the summer of 2008 the OECD stopped reporting part of the data in the online CRS table so that disaggregated commitments are now only displayed from 1995 onwards and disbursements from 2002. The reason is a concern over weak CRS coverage for earlier years. This weak coverage, however, is exactly what the data construction method developed in this chapter attempts to address. The CRS data I use was downloaded in December 2006, before this occurred. Data for earlier years can still be accessed in the CRS table by clicking on the 'related files' icon.

[36] In addition to the supply of experts, teachers and volunteers, and expenditure on research, equipment and materials, the DAC directive lists the cost of students and trainees, and the financing of development-oriented social and cultural programmes as part of TC (OECD, 2000a, pp. 59-62).

as agriculture, education, transportation, etc. Assistance is made available 'in cash' or 'in kind', with or without restriction on the specific use of the funds, but on the condition the recipient executes a development plan in favour of the sector concerned. Investment projects comprise schemes to increase and/or improve the recipient's stock of physical capital and financing the supply of goods and services in support of such schemes" (OECD, 2002, p. 22). This includes investment-related technical cooperation, which is "the financing of services by a donor country with the primary purpose of contributing to the design and/or implementation of a project or programme aiming to increase the physical capital stock of the recipient country. These services include consulting services, technical support, the provision of know-how linked to the execution of an investment project, and the contribution of the donor's own personnel to the actual implementation of the project (managers, technicians, skilled labour etc.)" (OECD, 2002, p. 22). Other (no mark) is the residual category.

Sector programme aid should for the most part be on-budget, as by definition programme aid involves a government to government transfer of resources. Technical cooperation, on the other hand, should be predominantly off-budget. The cost of providing training and scholarships in donor countries, remunerating experts and consultants, and financing equipment and administrative costs associated with TC mostly involves a direct payment from the donor government, rather than a transfer of money to the recipient government. Sundberg and Gelb (2006) argue technical assistance is often spent outside the recipient country, while Fagernäs and Roberts (2004a,b) and Feeny (2007) attribute discrepancies between donor and recipient reports of aid in the countries they are studying at least in part to the omission of technical assistance from the recipient governments' budgets. The extent to which investment projects and other aid are reported in recipient government budgets is less clear.

A very small part of education and health aid in CRS is listed under a combination of prefix codes (e.g. IP & TC). In these cases, I allocate an equal part of the aid amount to each of the prefix codes that make up the combination.

At this stage it is important to note that CRS does not record zeros. If no aid is given in a sector the observation is simply missing so in general it is difficult to tell whether an observation is missing because no aid is disbursed or because existing aid flows are not reported. Whenever total education or health disbursements are available, which is the case when at least one of the four prefix codes is available, I set missing values for the other prefix codes to zero. Similarly, whenever aggregate disbursements are available, missing observations for sectoral disbursements, as well as education and health prefix codes, are changed to zero. The prefix codes always sum to total education and health disbursements. Similarly, aggregate CRS disbursements equal the sum of the underlying sectors, apart from tiny discrepancies.[37] I also download CRS data on aggregate grants

[37]Throughout the data construction, tiny discrepancies between totals and their underlying components may arise, even if the former is (re)calculated explicitly as the sum of the latter. This is because Stata stores numbers as binary and many decimal numbers have no exact binary representation, which may lead to small calculation 'errors' (Cox, 2006; Gould, 2006). It would be possible to deal with this by transforming all variables into integers and then transforming them back after the data construction (Gould, 2006). I forego this

and loans, which will become useful later. Again, missing observations for these two variables are turned to zero whenever aggregate CRS disbursements are available. CRS grants and loans always sum to aggregate CRS disbursements.

The aggregate and sectoral disbursements thus obtained from CRS in a recipient-donor-year format form the backbone of the data construction. From here on I refer to these variables as CRS^{agg}_{RDY} and CRS^{s}_{RDY} (for $s = 1, \ldots, S$), respectively. CRS disbursements at the prefix code level are labelled $CRS^{s,p}_{RDY}$, where s now refers to the education or health sector and $p = IP, SP, TC, ONM$. Because these aid measures are incomplete I attempt to improve on them, which first of all requires data from DAC table 2a.

A.2 Development Assistance Committee table 2a

The data in DAC2a should be complete but does not allow to fully disaggregate aid according to sector or prefix code. I download data on grants and loans extended, again in a RDY format. Missing values for loans are set to zero when grants are observed, and vice versa. Total disbursements, $DAC2a^{agg}_{RDY}$, are then calculated as the sum of grants and loans. The OECD makes a distinction between Official Development Assistance (ODA) and Official Assistance (OA), where OA is simply ODA directed to countries on part II of the DAC list of aid recipients, comprised of transition countries and more advanced developing countries (OECD, 2000a, p. 11 and p. 64). Whether aid transferred to a given recipient is classified as ODA or OA may vary over time. While OECD (2002, p. 4) states that the CRS database contains both ODA and OA, in the CRS data I downloaded no observations are available for recipient-years that are listed on part II of the recipient list in DAC2a. As a result, I focus only on ODA in DAC2a and exclude part II recipient-years. Conversely, for Serbia CRS data is available but DAC2a data is not so Serbia is dropped from the sample.[38] In addition, I only select donors that are also available in DAC table 5, for reasons that will become clear shortly. Two donors are excluded from DAC2a because of this: GFATM (Global Fund to Fight Aids, Tuberculosis and Malaria) and UNFPA (United Nations Population Fund).

A.3 Calculating the amount of aid missing from CRS

I now have data on (supposedly) complete aggregate DAC2a disbursements and incomplete aggregate and sectoral CRS disbursements, both in a recipient-donor-year format. By subtracting CRS disbursements from DAC2a disbursements I obtain a residual for aggregate disbursements (RES^{agg}_{RDY}). For each RDY observation

option, because it adds another layer of complexity and because the discrepancies that arise are negligibly small. I do consistently store variables as 'double' in Stata, so as to keep discrepancies as small as possible.

[38]The dataset still contains 'Serbia & Montenegro, FRY' as a recipient for 1994-2004.

this residual captures the amount of disbursed aid that is missing from the CRS database:

$$RES^{agg}_{RDY} = DAC2a^{agg}_{RDY} - CRS^{agg}_{RDY} \tag{25}$$

The aim is to allocate this residual across sectors.

RES^{agg}_{RDY} is negative for quite a few observations. In the majority of such cases CRS disbursements exceed DAC2a disbursements by only a very small margin but there are also a number of observations where the difference is larger. I replace DAC2a grants (loans) by the CRS amount in all cases where CRS grants (loans) exceed DAC2a grants (loans). I then recalculate $DAC2a^{agg}_{RDY}$ as the sum of DAC2a grants and loans, and recalculate RES^{agg}_{RDY}. If the DAC2a value is negative and the CRS value is zero, however, no replacement is carried out, whereas if the DAC2a value is negative and the CRS value is non-zero the former *is* replaced by the latter.

The rationale for these adjustments is that it is very unlikely that aid is reported if it never actually took place. It is far more likely actual aid is underreported, i.e. it is more likely DAC2a figures are missing something when they are exceeded by CRS figures, even though they are supposed to be complete. It might also be the case that negative amounts of aid go unreported in CRS and this is what causes the CRS figures to exceed the DAC2a figures. This is less probable, however, since negative amounts of aid, which presumably capture the repayment of unused aid money or resources, are quite rare in the data.[39]

Applying this rationale consistently is also what leads me not to replace the DAC2a value by the CRS value if the former is negative and the latter is zero. A zero CRS value means no aid is reported to CRS, while the negative value for DAC2a implies there was some aid, albeit negative. The situation where DAC2a aid is negative and CRS aid is non-zero is more tricky. On the one hand, the DAC2a database is supposed to be complete so its value is more likely to be the true one but, on the other hand, negative amounts of aid are rare and it is difficult to interpret them, which tilts the balance of favour of the CRS figure. Hence, in this case I replace the negative DAC2a amount by the non-zero (and always positive) CRS amount. Because there are only a few such observations (9 for grants and 17 for loans, out of a total of 43216 RDY observations) this choice should not have a substantial impact on the data.

For some RDY observations (1230 in total) CRS data is available but DAC2a data is not.[40] For these observations no residual can be calculated. Even so, I do not delete these observations from the CRS database, they are simply treated as having a zero residual. Conversely, if observations are available in DAC2a but missing from CRS, all CRS variables are changed to zero so that the complete DAC2a value is recorded as a residual.

[39] Recall I am working with gross disbursements so these negative amounts of aid do not reflect loan repayments. In the RDY CRS data there is not a single negative observation for the aid variables I distinguish. In the DAC2a dataset 185 out of a total of 43216 RDY combinations are negative for grants and/or loans.

[40] Some of these observations arise because I have excluded donors GFATM and UNFPA from DAC2a, due to the fact that they are absent from DAC5.

Having calculated a total residual for each RDY observation, I collapse the dataset by summing over recipients, yielding a residual for aggregate disbursements in a donor-year format ($RES_{DY}^{agg,C}$, where DY stands for donor-year and C makes clear this residual is formed by *collapsing* RES_{RDY}^{agg} over all recipients):

$$RES_{DY}^{agg,C} = \sum_{R} RES_{RDY}^{agg} \tag{26}$$

While the RDY data contains 113 negative residuals, $RES_{DY}^{agg,C}$ is always positive. The reason for collapsing the dataset is that now, with data from one more DAC table, it becomes possible to allocate $RES_{DY}^{agg,C}$ across sectors for each donor-year.

A.4 Development Assistance Committee table 5

To do this, one more piece of information, which comes from DAC table 5, is needed. DAC5 comprises a sectoral disaggregation of total ODA but only in a donor-year format. While this means the data are not available from a recipient perspective, the advantage of DAC5 is that it should contain more complete information than CRS. I label total and sectoral aid from this table as $DAC5_{DY}^{agg}$ and $DAC5_{DY}^{s}$, respectively. As in CRS, the sectors of interest are: education (DAC5 sector code 110), health (120), commodity aid/general programme assistance (500), action relating to debt (600), donor administrative costs (910), support to NGOs (920) and other sectors (the sum of all remaining sector codes). Missing observations for sectoral aid are set to zero whenever $DAC5_{DY}^{agg}$ is available. A problem is that $DAC5_{DY}^{agg}$ is not always equal to the sum of the sectoral aid variables. Four observations show up with large discrepancies: AsDF (Asian Development Fund) 1996, AsDF 2002, France 1997, and IDB (Inter-American Development Bank) Special Fund 1996.

For AsDF 2002 and France 1997 $DAC5_{DY}^{agg}$ exceeds the sum of the sectoral aid variables. In both cases this is because the entry for total sector allocable aid exceeds the sum of its underlying series.[41] Hence, for both observations I scale up all sector allocable series so that their sum matches total sector allocable aid. This means education and health aid are scaled up but also the other sector allocable series, which make up part of other sector aid. Therefore, after scaling up, other sector aid is recalculated as the sum of the underlying sectors. For all other observations discrepancies are extremely small, most likely due to rounding errors. To get rid of these small discrepancies $DAC5_{DY}^{agg}$ is recalculated as the sum of the sectoral aid variables. For AsDF 1996 and IDB Special Fund 1996 the sectoral sum exceeds $DAC5_{DY}^{agg}$ so these observations are also taken care of in this way.

Lastly, from DAC5 I also download data that partition health and education aid into the four prefix codes,

[41] Sector allocable aid includes aid for social infrastructure and services (including education and health), economic infrastructure and services, production sectors, and multisector/crosscutting aid. What remains is aid that cannot be allocated across sectors: commodity aid/general programme assistance, action relating to debt, emergency assistance, administrative costs of donors, support to NGOs and unallocated/unspecified aid.

again in a donor-year format ($DAC5^{s,p}_{DY}$).[42] Because, for AsDF 2002 and France 1997, education and health aid have been scaled up (see previous paragraph) I also scale up the prefix codes for these observations so that they still sum to total education and health aid. As before, missing observations for the prefix codes are set to zero whenever at least one of the other prefix codes within the sector is observed.

Unfortunately, the prefix codes in DAC5 do not always sum to total education and health aid. There is one observation for which the education total exceeds the sum of the prefix codes, while for health there are three such observations. For these observations I scale up the prefix codes so that their sum matches the sector total. I then recalculate education and health totals as the sum of their prefix codes for all other observations. This takes care of the one observation in both sectors for which the sum of the prefix codes exceeds the sector total. It also sorts out the many observations for which there are extremely small discrepancies. As this leads to changes in the values of education and health aid I recalculate $DAC5^{agg}_{DY}$ as the sum of the underlying sectors to ensure consistency.

This means I now have, in donor-year format, (supposedly) complete aid data disaggregated by sector from DAC5 and incomplete aid data disaggregated by sector from the collapsed CRS dataset ($CRS^{agg}_{DY} = \sum_R CRS^{agg}_{RDY}$, $CRS^{s}_{DY} = \sum_R CRS^{s}_{RDY}$, $CRS^{s,p}_{DY} = \sum_R CRS^{s,p}_{RDY}$). The plan is to calculate sectoral residuals for each donor-year and to use these to allocate each donor's total residual across sectors in each year. Going back to the data in recipient-donor-year format (RES^{agg}_{RDY}) this donor- and year-specific sectoral allocation of the total residual is then applied to all recipients that receive aid from the relevant donor in a given year that is not accounted for in CRS.

There is, however, one problem that needs to be solved before proceeding. The sectoral residuals must be calculated from DAC5 data, whereas the total residual is based on DAC2a data (as DAC5 is not available in RDY format). Apart from the possibility of reporting inconsistencies between the two tables, a bigger problem arises because donors have a choice in DAC5 to report either commitments or disbursements. I received some information from the DAC for the years 2001-2004 as to who reports what. Out of the 127 DY observations with data in DAC5 for which I have this information 72 refer to disbursements and 55 to commitments. However, these 55 observations include many of the larger donors, such as the United States, Japan, the European Commission, Germany and France.

As a consequence I scale all DAC5 aid variables, including the education and health prefix codes, by the ratio of aggregate DAC2a disbursements to total DAC5 ODA so that the sectoral aid variables from DAC5 sum

[42] In contrast with the CRS database, DAC5 classifies combinations of prefix codes as ONM (OECD, 2000a, p. 118). My decision to instead allocate an equal part of the aid amount to each of the prefix codes that make up the combination in the CRS data should have little effect, though, since only a very small part of education and health aid is listed under a combination of prefix codes.

to DAC2a aggregate disbursements:

$$\widehat{DAC5}^{s}_{DY} = DAC2a^{agg}_{DY}\left(\frac{DAC5^{s}_{DY}}{DAC5^{agg}_{DY}}\right) \tag{27}$$

for $s = 1, \ldots, S$, and:

$$\widehat{DAC5}^{s,p}_{DY} = DAC2a^{agg}_{DY}\left(\frac{DAC5^{s,p}_{DY}}{DAC5^{agg}_{DY}}\right) \tag{28}$$

for the education and health sectors and $p = IP, SP, TC, ONM$. This amounts to assuming that the sectoral allocation in DAC5 (of commitments or disbursements) is an accurate guide to the sectoral allocation of DAC2a disbursements. The correlation between $DAC5^{agg}_{DY}$ and $DAC2a^{agg}_{DY}$, at least, is very high (0.90). A few positive $DAC5^{agg}_{DY}$ values are scaled to zero because $DAC2a^{agg}_{DY}$ is zero but since these observations have no aggregate disbursements residual that needs to be allocated anyway this is not a problem. Scaling the data in this way ensures that the sectoral aid variables from DAC5 sum to DAC2a aggregate disbursements. This allows for a calculation of sectoral residuals that is more consistent with the calculation of the total residual, which is based on $DAC2a^{agg}_{RDY}$.

If, after the scaling, sectoral values in CRS exceed those in DAC5, I replace the latter by the former. I first carry out this replacement at the level of the prefix codes and recalculate total education and health aid as the sum of their prefix codes. I then repeat this strategy for all sectoral aid variables. At this stage the only changes for education and health aid occur for observations for which there is no prefix code disaggregation. So, after these changes the prefix codes still sum to total education and health aid for all observations that have data on both. As before, the DAC5 value is not replaced by the CRS value if the DAC5 value is negative and the CRS value is zero. However, if the DAC5 value is negative and the CRS value is non-zero then the former is replaced by the latter. The adjustments are limited in number and size, which is brought out by the high correlation (0.99) between the sum of the DAC5 sectoral aid variables (after scaling and replacement: $\sum_{s=1}^{S}\widehat{DAC5}^{s}_{DY}$) and $DAC2a^{agg}_{DY}$.

A.5 Allocating the residual across sectors

The total residual in donor-year format, RES_{DY}, is now calculated as the sum of the DAC5 sectoral aid variables minus aggregate CRS disbursements:

$$RES^{agg}_{DY} = \sum_{s=1}^{S}\widehat{DAC5}^{s}_{DY} - CRS^{agg}_{DY} \tag{29}$$

The correlation with the collapsed residual that was computed earlier from the recipient-donor-year dataset ($RES^{agg,C}_{DY}$) is 0.97. Sectoral (prefix) residuals in this DY format are calculated as the difference between

DAC5 sectoral (prefix) aid variables and sectoral (prefix) CRS disbursements:

$$RES^{s}_{DY} = \widehat{DAC5}^{s}_{DY} - CRS^{s}_{DY} \tag{30}$$

$$RES^{s,p}_{DY} = \widehat{DAC5}^{s,p}_{DY} - CRS^{s,p}_{DY} \tag{31}$$

The sectoral residuals sum to RES^{agg}_{DY} and residuals for the prefix codes sum to the total residuals for education and health. Whenever the CRS value is missing, the full DAC5 value is recorded as residual, as before.

Two sectoral residuals are negative. Finland 1991 has a negative residual for health IP (the DAC5 value is negative, while the CRS value is zero). For this observation I turn the health prefix code residuals to missing. UK 1996 has a negative residual for action relating to debt. Because this observation has a large total residual it would be a shame to lose it. Moreover, the absolute value of the negative action relating to debt residual is less than 0.1% of the total residual. Therefore, I set the action relating to debt residual to zero for this observation and recalculate RES^{agg}_{DY} as the sum of the sectoral residuals.

Now it is possible to calculate the shares of the sector residuals in the total residual ($SHRES^{s}_{DY}$), as well as the share of the prefix code residuals in the total education and health residuals ($SHRES^{s,p}_{DY}$):

$$SHRES^{s}_{DY} = \frac{RES^{s}_{DY}}{\sum_{s=1}^{S} RES^{s}_{DY}} \tag{32}$$

$$SHRES^{s,p}_{DY} = \frac{RES^{s,p}_{DY}}{\sum_{p} RES^{s,p}_{DY}} \tag{33}$$

This donor- and year-specific allocation of RES^{agg}_{DY} across sectors is then applied to the total residual calculated in the original RDY format (RES^{agg}_{RDY}):

$$\widehat{RES}^{s}_{RDY} = SHRES^{s}_{DY} RES^{agg}_{RDY} \tag{34}$$

That is, I apply the sectoral residual shares of a given donor-year to the total residuals of all recipients to which the donor gives aid in that year that is not fully accounted for in CRS. In other words, I assume the sectoral allocation of a donor's total residual is the same for all recipients with which this donor has a residual. For instance, if Botswana and Tanzania receive an unallocated residual from the US in 2004, and (32) shows that half of the total residual of the US in 2004 consists of education aid and half consists of health aid, then for both Botswana and Tanzania half of the total residual with the US in 2004 is classified as education aid and half as health aid. Total education and health residuals are allocated across prefix codes in the same way:

$$\widehat{RES}^{s,p}_{RDY} = SHRES^{s,p}_{DY} \widehat{RES}^{s}_{RDY} \tag{35}$$

A.6 Creating more complete sectoral aid disbursements

I add the sectoral residuals to the CRS disbursements in the RDY database, and likewise for the education and health prefix codes:

$$\widetilde{CRS}^{s}_{RDY} = CRS^{s}_{RDY} + \widehat{RES}^{s}_{RDY} \tag{36}$$
$$\widetilde{CRS}^{s,p}_{RDY} = CRS^{s,p}_{RDY} + \widehat{RES}^{s,p}_{RDY} \tag{37}$$

For some observations insufficient information is available in DAC5 to allocate RES^{agg}_{RDY} across sectors.[43] As a result, the sum of the newly calculated sectoral variables does not necessarily equal $DAC2a^{agg}_{RDY}$.[44] Similarly, education and health prefix codes do not always sum to the education and health total because for some donors insufficient information is available to allocate the education and health residuals across prefix codes.

Therefore, as a final step in the data construction, after collapsing the data to a recipient-year (RY) format, I scale the sectoral disbursements so that their sum equals a plausible measure of aggregate disbursements received. Before collapsing the data I replace missing $DAC2a^{agg}_{RDY}$ by CRS^{agg}_{RDY} for the 1230 RDY observations that have CRS data but are missing from DAC2a (see above).[45]

I collapse the RDY dataset by summing over donors:

$$\widetilde{CRS}^{s}_{RY} = \sum_{D} \widetilde{CRS}^{s}_{RDY} \tag{38}$$
$$\widetilde{CRS}^{s,p}_{RY} = \sum_{D} \widetilde{CRS}^{s,p}_{RDY} \tag{39}$$

In this final recipient-year (RY) dataset there are observations for which both aggregate DAC2a and CRS disbursements are zero. The reason why these observations are zero rather than missing (as one would expect) is that Stata turns missing values to zero when collapsing data. I turn all aid variables to missing for these observations. In addition, there are seven observations with non-zero aggregate DAC2a disbursements but zeros for all sectoral aid variables. Since, for these observations, there is no information at all about the allocation of aggregate disbursements across sectors, all variables are turned to missing. Similarly, there is one observation

[43] While bilateral donors' ODA is typically available for all years in DAC5, data for multilateral donors is more patchy. Data for IBRD (International Bank for Reconstruction and Development) and IDA (International Development Association), for instance, is only available for 4 and 5 years in the beginning of the 90s, respectively (IBRD is also missing from DAC2a). In the years with data, the magnitude of aggregate ODA in DAC5 and aggregate disbursements in DAC2a is relatively similar for both bilateral and multilateral donors (also see below). Generally speaking, multilateral donors' coverage is worse in CRS as well.

[44] Conversely, there are also observations where aggregate CRS disbursements are zero but a DAC2a total is available that has been allocated across the different sectors. For these recipients with no sectoral CRS data the sectoral disbursements I end up with are based entirely on how the residuals of the donors that deal with this recipient are allocated across sectors.

[45] Some of these 1230 observations involve the two donors (GFATM and UNFPA) that are available in DAC2a but missing from DAC5. With hindsight, I should not have excluded these donors from DAC2a. In fact, I could have included all available donors in DAC2a even if they are absent from DAC5 or CRS and then sum over all donors to obtain aggregate disbursements in a RY format. Before I scale the constructed sectoral disbursements to a plausible measure of aggregate disbursements, however, I also download RY data from DAC2a with 'all donors' as donors, and use this variable as a candidate measure of aggregate disbursements received (see below), so the effect of this omission – if any – should be extremely small.

with zeros for all health prefix codes, but a non-zero health total. For this observation health prefix codes are changed to missing.

As before the collapse, when I sum the sectoral disbursements I do not always get a number that equals aggregate DAC2a disbursements ($DAC2a^{agg}_{RY} = \sum_D DAC2a^{agg}_{RDY}$), and, similarly, the sum of the prefix codes does not always equal total education and health aid. I first scale the prefix codes so that their sum equals total education and health aid. This is done by multiplying each prefix code with the ratio of total sectoral (education or health) aid to the sum of the prefix codes:

$$\overline{CRS}^{s,p}_{RY} = \widetilde{CRS}^{s}_{RY} \left(\frac{\widetilde{CRS}^{s,p}_{RY}}{\sum_p \widetilde{CRS}^{s,p}_{RY}} \right) \tag{40}$$

For Chinese Taipei (more commonly known as Taiwan) several years have negative values for total health aid while the sum of the health prefix codes is positive. In addition, in the remaining observed years (except 1990) the sum of the health prefix codes always exceeds total health aid and these are the only observations in the dataset for which this is the case. Similarly, in all observed years except 1990 Chinese Taipei has a value for total education aid that is smaller than the sum of the prefix codes (the latter is also the case for Somalia 1997). This seems to suggest data for Chinese Taipei contains a great deal of measurement error. Given that Chinese Taipei has no data after 1996 in any case, it is dropped from the dataset in its entirety. For both sectors Cayman islands 1991 has a negative prefix sum. However, because total education and health aid are also negative, scaling should not be a problem for this observation. For now, I keep this observation and simply apply the scaling, as it will be dropped at a later stage for other reasons in any case.

I now apply the same strategy to the sectoral aid variables to make sure their sum matches an aggregate measure of disbursements received. Recall that aggregate DAC2a disbursements in this RY format are calculated by summing DAC2a disbursements in the RDY format over all donors, and that donors that are missing from DAC5 or CRS were not selected when downloading data for $DAC2a^{agg}_{RDY}$. Consequently, aid from these donors is not included in $DAC2a^{agg}_{RY}$. Therefore, in addition to $DAC2a^{agg}_{RY}$, I download grants and loans from DAC2a in a RY format, selecting ‘all donors (total)’ in the donor dimension. Missing grants are set to zero when loans are observed, and vice versa. Total disbursements, $DAC2a^{agg}_{RY,AD}$ (AD stands for all donors), are calculated as the sum of grants and loans extended. The correlation between this measure and $DAC2a^{agg}_{RY}$ is extremely high (0.99). The sum of the sectoral variables has a similarly high correlation with both measures.

I scale the sectoral variables so that their sum equals the maximum of $DAC2a^{agg}_{RY}$ and $DAC2a^{agg}_{RY,AD}$. Again, this follows the rationale that it is unlikely non-existing aid is reported, so the higher figure should be the most accurate one. While $DAC2a^{agg}_{RY,AD}$ should include aid from more donors, $DAC2a^{agg}_{RDY}$ (on which $DAC2a^{agg}_{RY}$ is based) has been adjusted upwards for those observations where it is exceeded by aggregate CRS disbursements (see above).

For 4 observations (Costa Rica 1992, Mexico 1992, Panama 1992, Saudi Arabia 1991) the sum of the sectoral aid variables ($\sum_{s=1}^{S} \widetilde{CRS}^{s}_{RY}$) slightly exceeds $DAC2a^{agg}_{RY}$ (for some other observations the difference is negligibly small and due to the way Stata stores data). This may arise if a recipient receives a negative total residual from a donor for which no sectoral allocation can be calculated. Since $DAC2a^{agg}_{RY}$ incorporates this negative amount of aid while the sectoral aid variables do not, the sectoral sum may exceed $DAC2a^{agg}_{RY}$ if the negative residual is not offset by positive residuals from other donors for which the sectoral allocation is also lacking. For these observations $\sum_{s=1}^{S} \widetilde{CRS}^{s}_{RY}$ may also exceed $DAC2a^{agg}_{RY,AD}$, which here is only the case for Panama 1992. Since $\sum_{s=1}^{S} \widetilde{CRS}^{s}_{RY}$ only exceeds $DAC2a^{agg}_{RY,AD}$ and $DAC2a^{agg}_{RY}$ if it does not incorporate negative amounts of aid that are known to have taken place but that I was not able to allocate across sectors, it is likely to exaggerate aid disbursements for the observations where this is the case. As a result, I scale only to the maximum of $DAC2a^{agg}_{RY,AD}$ and $DAC2a^{agg}_{RY}$. This maximum value is labelled $DISB_{RY}$. Consequently, the final measures of sectoral and prefix code aid disbursements are:

$$\widehat{CRS}^{s}_{RY} = DISB_{RY} \left(\frac{\widetilde{CRS}^{s}_{RY}}{\sum_{s} \widetilde{CRS}^{s}_{RY}} \right) \tag{41}$$

$$\widehat{CRS}^{s,p}_{RY} = DISB_{RY} \left(\frac{\overline{CRS}^{s,p}_{RY}}{\sum_{s} \widetilde{CRS}^{s}_{RY}} \right) \tag{42}$$

One observation (Cayman islands 1991) has a negative sectoral sum. For this observation the only residual that can be allocated across sectors is negative, whereas for the two donors with a positive residual no sectoral allocation is available. Hence, each sectoral aid variable, and their sum, is negative, whereas $DAC2a^{agg}_{RY}$ is positive. I turn all variables to missing for this observation.

There are ten recipient-year observations with sectoral CRS data but missing DAC2a data. When examining the time series around these observations in more detail, for all but one (Slovenia 1992) it is evident that aggregate CRS disbursements are a lot lower than aggregate DAC2a disbursements in subsequent years. Hence, I choose not to rely solely on the CRS data, which could seriously underestimate the total amount of aid, and instead turn all variables to missing when $DAC2a^{agg}_{RY,AD}$ is missing.

Finally, I drop high-income countries, defined as countries with a 2005 GNI per capita of 10726 US$ or more (following World Bank, 2006c). Many of the high-income countries are small islands (e.g. Antigua and Barbuda, Aruba, Netherlands Antilles) or oil exporters (e.g. Kuwait, Qatar, United Arab Emirates). Two remaining observations (Turkey 2000 for education SP aid and Barbados 2001 for health SP aid) are smaller than zero. Since in both cases it concerns extremely small negative values (less than 0.0001% of GDP in absolute value) and since negative aid values are difficult to interpret, I set these observations to zero.

Table 11 shows summary statistics for the scaling that takes place in the final step of the data construction (see equations (41) and (42)). $scaling$ is computed as the ratio of the sum of the constructed sectoral disbursements (before scaling) to $DISB_{RY}$:[46]

$$scaling = \frac{\sum_{s=1}^{S} \widetilde{CRS}_{RY}^{s}}{DISB_{RY}} \tag{43}$$

This is compared to the scaling that would take place if I simply scale sectoral CRS disbursements so that their sum matches a measure of total aggregate disbursements, following the logic behind equation (11):

$$scaling_{CRS} = \frac{\sum_{s=1}^{S} CRS_{RY}^{s}}{DISB_{RY}} \tag{44}$$

As can be seen from table 11, the difference between $scaling$ and $scaling_{CRS}$ is large. On average, the constructed disbursements before scaling make up more than 76% of aggregate, complete disbursements, whereas for CRS disbursements this is only 31.9%. This difference reflects the information added to the sectoral CRS disbursements by the data construction method described in this appendix. For the majority of observations the scaling performed in the final step of the data construction is limited in magnitude and a lot smaller than if CRS sectoral disbursements are scaled without any adjustment. For instance, for more than three quarters of observations CRS disbursements constitute less than half of aggregate aid. For the constructed sectoral disbursements this is the case for less than 10% of observations. This makes it more likely that the sectoral allocation of the aid data before scaling is a reasonable reflection of the actual sectoral allocation one would find if data were complete. This is again the best that can be done with the available data, and not scaling the sectoral disbursements runs the risk of underestimating the amount of aid received.

Table 11: Scaling variables

	$scaling$	$scaling_{CRS}$
Observations	2192	2192
Mean	0.768	0.319
Standard deviation	0.191	0.264
Minimum	0.016	0
1st percentile	0.174	0
5th percentile	0.391	0
10th percentile	0.515	0.015
25th percentile	0.656	0.097
Median	0.804	0.258
75th percentile	0.925	0.494
90th percentile	0.981	0.726
95th percentile	0.996	0.843
99th percentile	1	0.981
Maximum	1.128	1

[46]Note the maximum value exceeds one. This is the observation for Panama 1992.

References

ADAM, C. S., P.-Å. ANDERSSON, A. BIGSTEN, P. COLLIER, AND S. O'CONNELL (1994): "Evaluation of Swedish development co-operation with Zambia," Ministry of Foreign Affairs, Secretariat for analysis of Swedish development assistance, Report 6.

ANDERSON, T., AND C. HSIAO (1981): "Estimation of dynamic models with error components," *Journal of the American Statistical Association*, 76(375), 598–606.

——— (1982): "Formulation and estimation of dynamic models using panel data," *Journal of Econometrics*, 18(1), 47–82.

ARELLANO, M. (1987): "Computing robust standard errors for within-groups estimators," *Oxford Bulletin of Economics and Statistics*, 49(4), 431–434.

——— (1993): "On the testing of correlated effects with panel data," *Journal of Econometrics*, 59(1-2), 87–97.

ARELLANO, M., AND S. BOND (1991): "Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations," *Review of Economic Studies*, 58(2), 277–297.

ARELLANO, M., AND O. BOVER (1995): "Another look at the instrumental variable estimation of error-components models," *Journal of Econometrics*, 68(1), 29–51.

ASIEDU, E., AND B. NANDWA (2007): "On the impact of foreign aid in education on growth: how relevant is the heterogeneity of aid flows and the heterogeneity of aid recipients," *Review of World Economics*, 143(4), 631–649.

AZAM, J.-P., AND J.-J. LAFFONT (2003): "Contracting for aid," *Journal of Development Economics*, 70(1), 25–58.

BAQIR, R. (2002): "Social sector spending in a panel of countries," International Monetary Fund, Working Paper 02/35.

BASER, H., AND P. MORGAN (2001): "The pooling of technical assistance: an overview based on field research experience in six African countries," a study for the Directorate-General for Development Cooperation (DGIS), Ministry of Foreign Affairs of the Netherlands.

BELSLEY, D. A., E. KUH, AND R. E. WELSCH (1980): *Regression diagnostics: identifying influential data and sources of collinearity*. John Wiley & Sons, New York, United States, 292 p.

BERTRAND, M., E. DUFLO, AND S. MULLAINATHAN (2004): "How much should we trust differences-in-differences estimates," *Quarterly Journal of Economics*, 119(1), 249–275.

BLUNDELL, R., AND S. BOND (1998): “Initial conditions and moment restrictions in dynamic panel data models,” *Journal of Econometrics*, 87(1), 115–143.

BOLLEN, K. A., AND R. W. JACKMAN (1990): “Regression diagnostics: an expository treatment of outliers and influential cases,” in *Modern methods of data analysis*, ed. by J. Fox, and J. S. Long, pp. 257–291. Sage, Newbury Park, London.

BOND, S. R. (2002): “Dynamic panel data models: a guide to micro data methods and practice,” *Portuguese Economic Journal*, 1(2), 141–162.

CHATTERJEE, S., P. GIULIANO, AND I. KAYA (2007): “Where has all the money gone? Foreign aid and the quest for growth,” Institute for the Study of Labor (IZA), Bonn, Germany, Discussion Paper No. 2858.

CLEMENS, M. A., S. RADELET, AND R. BHAVNANI (2004): “Counting chickens when they hatch: the short-term effects of aid on growth,” Center for Global Development, Working Paper Number 44 (revised edition, December 2, 2004).

COX, N. J. (2006): “Stata tip 33: sweet sixteen: hexadecimal formats and precision problems,” *Stata Journal*, 6(2), 282–283.

DEPETRIS CHAUVIN, N., AND A. KRAAY (2005): “What has 100 billion dollars worth of debt relief done for low-income countries,” World Bank, mimeo.

DEVARAJAN, S., A. S. RAJKUMAR, AND V. SWAROOP (2007): “What does aid to Africa finance,” in *Theory and practice of foreign aid*, ed. by S. Lahiri. Elsevier, Amsterdam, The Netherlands, 438 p.

DIKHANOV, Y. (2004): “Historical PV of debt in developing countries, 1980-2002,” World Bank, mimeo.

DREHER, A. (2006): “The influence of globalization on taxes and social policy: an empirical analysis for OECD countries,” *European Journal of Political Economy*, 22(1), 179–201.

DREHER, A., P. NUNNENKAMP, AND R. THIELE (2008): “Does aid for education educate children? Evidence from panel data,” *World Bank Economic Review*, 22(2), 291–314.

DREHER, A., J.-E. STURM, AND H. W. URSPRUNG (2006): “The impact of globalization on the composition of government expenditures: evidence from panel data,” University of Munich, Center for Economic Studies Information and Forschung (CESifo), Working Paper No. 1755.

DRUKKER, D. M. (2003): “Testing for serial correlation in linear panel-data models,” *Stata Journal*, 3(2), 168–177.

FAGERNÄS, S., AND J. ROBERTS (2004a): "The fiscal effects of aid in Uganda," Overseas Development Institute (ODI), Economics and Statistics Analysis Unit, Working Paper 9.

——— (2004b): "The fiscal effects of aid in Zambia," Overseas Development Institute (ODI), Economics and Statistics Analysis Unit, Working Paper 10.

FAGERNÄS, S., AND C. SCHURICH (2004): "The fiscal effects of aid in Malawi," Overseas Development Institute (ODI), Economics and Statistics Analysis Unit, Working Paper 7.

FEENY, S. (2007): "Foreign aid and fiscal governance in Melanesia," *World Development*, 35(3), 439–453.

FEYZIOGLU, T., V. SWAROOP, AND M. ZHU (1998): "A panel data analysis of the fungibility of foreign aid," *World Bank Economic Review*, 12(1), 29–58.

FRANCO-RODRIGUEZ, S. (2000): "Recent developments in fiscal response with an application to Costa Rica," *Journal of International Development*, 12(3), 429–441.

FRANCO-RODRIGUEZ, S., O. MORRISSEY, AND M. MCGILLIVRAY (1998): "Aid and the public sector in Pakistan: evidence with endogenous aid," *World Development*, 26(7), 1241–1250.

GANG, I. N., AND H. A. KHAN (1991): "Foreign aid, taxes, and public investment," *Journal of Development Economics*, 34(1-2), 355–369.

GERDTHAM, U.-G., AND B. JÖNSSON (2000): "International comparisons of health expenditure: theory, data, and econometric analysis," in *Handbook of health economics*, ed. by A. J. Culyer, and J. P. Newhouse, pp. 11–53. Elsevier Science, Amsterdam.

GOULD, W. (2006): "Mata matters: precision," *Stata Journal*, 6(4), 550–560.

GRAMLICH, E. M. (1977): "Intergovernmental grants: a review of the empirical literature," in *The political economy of fiscal federalism*, ed. by W. E. Oates, pp. 219–239. Lexington Books, Lexington, Massachusetts, United States.

GUPTA, S., B. CLEMENTS, AND E. TIONGSON (1998): "Public spending on human development," *Finance and Development*, 35(3), 10–13.

GUPTA, S., L. DICKS-MIREAUX, R. KHEMANI, C. MCDONALD, AND M. VERHOEVEN (2000): "Social issues in IMF-supported programs," International Monetary Fund, Occasional Paper 191.

HANSEN, L. P. (1982): "Large sample properties of generalized method of moments estimators," *Econometrica*, 50(4), 1029–1054.

HELLER, P. S. (1975): "A model of public fiscal behavior in developing countries: aid, investment and taxation," *American Economic Review*, 65(3), 429–445.

HEPP, R. (2005): "Health expenditures under the HIPC initiative," University of California, Davis, Department of Economics, mimeo.

HOLTZ-EAKIN, D., W. NEWEY, AND H. S. ROSEN (1988): "Estimating vector autoregressions with panel data," *Econometrica*, 56(6), 1371–1395.

IDD AND ASSOCIATES (2006): "Evaluation of general budget support: synthesis report," International Development Department, School of Public Policy, University of Birmingham.

INTERNATIONAL MONETARY FUND (2006): "Government Finance Statistics," June 2006 CD-ROM.

KASUGA, H. (2007): "The Millennium Development Goals and aid allocation: which donors give high-quality aid," Research Institute of Economy, Trade and Industry (RIETI), Tokyo, Japan, Discussion Paper 07-E-050.

KAUFMAN, R. R., AND A. SEGURA-UBIERGO (2001): "Globalization, domestic politics, and social spending in Latin America: a time-series cross-section analysis, 1973-97," *World Politics*, 53(4), 553–587.

KÉZDI, G. (2004): "Robust standard error estimation in fixed-effects panel models," *Hungarian Statistical Review*, Special Number 9, 95–116.

KHILJI, N. M., AND E. M. ZAMPELLI (1991): "The fungibility of US assistance to developing countries and the impact on recipient expenditures: a case study of Pakistan," *World Development*, 19(8), 1095–1105.

——— (1994): "The fungibility of US military and non-military assistance and the impacts on expenditures of major aid recipients," *Journal of Development Economics*, 43(2), 345–362.

LAPORTE, A., AND F. WINDMEIJER (2005): "Estimation of panel data models with binary indicators when treatment effects are not constant over time," *Economics Letters*, 88(3), 389–396.

LORA, E., AND M. OLIVERA (2007): "Public debt and social expenditure: friends or foes?," *Emerging Market Review*, 8(4), 299–310.

LU, C., M. T. SCHNEIDER, P. GUBBINS, K. LEACH-KEMON, D. JAMISON, AND C. J. MURRAY (2010): "Public financing of health in developing countries: a cross-national systematic analysis," *Lancet*, 375(9723), 1375–1387.

MARSHALL, M. G., AND K. JAGGERS (2007): "Polity IV project: dataset users' manual," Center for Systemic Peace, mimeo.

MAVROTAS, G. (2002a): "Aid and growth in India: some evidence from disaggregated aid data," *South Asia Economic Journal*, 3(1), 19–49.

——— (2002b): "Foreign aid and fiscal response: does aid disaggregation matter," *Weltwirtschaftliches Archiv*, 138(3), 534–559.

——— (2005): "Aid heterogeneity: looking at aid effectiveness from a different angle," *Journal of International Development*, 17(8), 1019–1036.

MAVROTAS, G., AND B. OUATTARA (2006): "Aid disaggregation and the public sector in aid-recipient economies: some evidence from Côte d'Ivoire," *Review of Development Economics*, 10(3), 434–451.

MCGILLIVRAY, M., AND O. MORRISSEY (2004): "Fiscal effects of aid," in *Fiscal policy for development: poverty, reconstruction and growth*, ed. by T. Addison, and A. Roe, pp. 72–96. Palgrave Macmillan, Basingstoke, Hampshire, United Kingdom.

MCGILLIVRAY, M., AND O. MORRISSEY (2000): "Aid fungibility in Assessing Aid: red herring or true concern," *Journal of International Development*, 12(3), 413–428.

——— (2001): "Aid illusion and public sector fiscal behaviour," *Journal of Development Studies*, 37(6), 118–136.

MCGILLIVRAY, M., AND B. OUATTARA (2005): "Aid, debt burden and government fiscal behaviour in Côte d'Ivoire," *Journal of African Economies*, 14(2), 247–269.

MCGUIRE, M. (1982): "US assistance, Israeli allocation, and the arms race in the Middle East: an analysis of three interdependent resource allocation processes," *Journal of Conflict Resolution*, 26(2), 199–235.

——— (1987): "Foreign assistance, investment, and defense: a methodological study with an application to Israel, 1960-1979," *Economic Development and Cultural Change*, 35(4), 847–873.

MICHAELOWA, K., AND A. WEBER (2006): "Aid effectiveness reconsidered: panel data evidence for the education sector," Hamburg Institute for International Economics (HWWA), Discussion Paper 264 (revised version, 2006).

MISHRA, P., AND D. NEWHOUSE (2007): "Health aid and infant mortality," International Monetary Fund, Working Paper 07/100.

NICKELL, S. (1981): "Biases in dynamic models with fixed effects," *Econometrica*, 49(6), 1417–1426.

OECD (2000a): "DAC Statistical Reporting Directives," .

——— (2000b): "Handbook for reporting debt reorganisation on the DAC questionnaire," .

——— (2002): "Reporting Directives for the Creditor Reporting System," .

——— (2007a): "DAC Statistical Reporting Directives," .

——— (2007b): "Reporting Directives for the Creditor Reporting System," .

OSEI, R., O. MORRISSEY, AND T. LLOYD (2005): "The fiscal effects of aid in Ghana," *Journal of International Development*, 17(8), 1037–1053.

OUATTARA, B. (2006): "Aid, debt and fiscal policies in Senegal," *Journal of International Development*, 18(8), 1105–1122.

——— (2007): "Foreign aid, public savings displacement and aid dependency in Côte d'Ivoire: an aid disaggregation approach," *Oxford Development Studies*, 35(1), 33–46.

PACK, H., AND J. R. PACK (1990): "Is foreign aid fungible? The case of Indonesia," *Economic Journal*, 100(399), 188–194.

——— (1993): "Foreign aid and the question of fungibility," *Review of Economics and Statistics*, 75(2), 258–265.

——— (1999): "Foreign aid and fiscal stress," in *Development, duality, and the international economic regime: essays in honor of Gustav Ranis*, ed. by G. R. Saxonhouse, and T. Srinivasan, pp. 452–476. University of Michigan Press, Ann Arbor, Michigan, United States.

PETTERSSON, J. (2007a): "Child mortality: is aid fungibility in pro-poor expenditure sectors decisive," *Review of World Economics*, 143(4), 673–693.

——— (2007b): "Foreign sectoral aid fungibility, growth and poverty reduction," *Journal of International Development*, 19(8), 1074–1098.

RAVISHANKAR, N., P. GUBBINS, R. J. COOLEY, K. LEACH-KEMON, C. M. MICHAUD, D. T. JAMISON, AND C. J. MURRAY (2009): "Financing of global health: tracking development assistance for health from 1990 to 2007," *Lancet*, 373(9681), 2113–2124.

REPUBLIC OF LIBERIA MINISTRY OF FINANCE (2009): "Draft national budget for the fiscal year July 1, 2009 – June 30, 2010," .

RODRIK, D. (1998): "Why do more open economies have bigger governments," *Journal of Political Economy*, 106(5), 997–1032.

ROODMAN, D. (2008): "Through the looking glass, and what OLS found there: on growth, foreign aid, and reverse causality," Center for Global Development, Working Paper Number 137.

——— (2009a): "How to do xtabond2: an introduction to difference and system GMM in Stata," *Stata Journal*, 9(1), 86–136.

——— (2009b): "A note on the theme of too many instruments," *Oxford Bulletin of Economics and Statistics*, 71(1), 135–158.

SARGAN, J. (1958): "The estimation of economic relationships using instrumental variables," *Econometrica*, 26(3), 393–415.

SCHAFFER, M., AND S. STILLMAN (2006): "Stata module to calculate tests of overidentifying restrictions after xtreg, xtivreg, xtivreg2 and xthtaylor," http://ideas.repec.org/c/boc/bocode/s456779.html.

SPANOS, A., AND A. MCGUIRK (2002): "The problem of near-multicollinearity revisited: erratic vs systematic volatility," *Journal of Econometrics*, 108(2), 365–393.

STOCK, J. H., AND M. W. WATSON (2008): "Heteroscedasticity-robust standard errors for fixed effects panel data regression," *Econometrica*, 76(1), 155–174.

SUNDBERG, M., AND A. GELB (2006): "Making aid work," *Finance and Development*, 43(4), 14–17.

SWAROOP, V., S. JHA, AND A. S. RAJKUMAR (2000): "Fiscal effects of foreign aid in a federal system of governance: the case of India," *Journal of Public Economics*, 77(3), 307–330.

TEMPLE, J. (2000): "Growth regressions and what the textbooks don't tell you," *Bulletin of Economic Research*, 52(3), 181–205.

THE POLITICAL RISK SERVICES GROUP (2008): "International Country Risk Guide methodology," http://www.prsgroup.com/PDFS/icrgmethodology.pdf (consulted 7 July 2008).

THIELE, R., P. NUNNENKAMP, AND A. DREHER (2007): "Do donors target aid in line with the Millennium Development Goals? A sector perspective of aid allocation," *Review of World Economics*, 143(4), 596–630.

THOMAS, A. (2006): "Do debt-service saving and grants boost social expenditures," International Monetary Fund, Working Paper 06/180.

UNITED NATIONS (2006): "The Millennium Development Goals report 2006," United Nations, New York, United States.

UNITED NATIONS GENERAL ASSEMBLY (2000): “Resolution adopted by the General Assembly 55/2. United Nations Millennium Declaration,” United Nations, New York, United States.

VAN DE SIJPE, N. (2010): “Foreign aid and government behaviour,” Thesis (DPhil), University of Oxford, Department of Economics.

WHITE, H. (1980): “A heteroscedasticity-consistent covariance matrix estimator and a direct test for heteroscedasticity,” *Econometrica*, 48(4), 817–838.

WILLIAMSON, C. R. (2008): “Foreign aid and human development: the impact of foreign aid to the health sector,” *Southern Economic Journal*, 75(1), 188–207.

WINDMEIJER, F. (2005): “A finite sample correction for the variance of linear efficient two-step GMM estimators,” *Journal of Econometrics*, 126(1), 25–51.

WOLF, S. (2007): “Does aid improve public service delivery,” *Review of World Economics*, 143(4), 650–672.

WOOLDRIDGE, J. M. (2002): *Econometric analysis of cross section and panel data*. MIT Press, Cambridge, Massachusetts, United States, 752 p.

WORLD BANK (2006a): “Edstats,” .

——— (2006b): “Global Development Finance,” .

——— (2006c): “World Development Indicators,” .