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

Nicolas Van de Sijpe,
Centre for the Study of African Economies (CSAE),
Department of Economics,
University of Oxford
nicolas.vandesijpe@economics.ox.ac.uk

May 31, 2007

PRELIMINARY DRAFT, PLEASE DO NOT QUOTE.

Abstract

This paper challenges the prevailing view among donors and development scholars that foreign aid is fungible. Unlike the bulk of previous empirical studies, I employ panel data that contains information on the specific purpose for which aid is given. This allows me to link aid given for education and health 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. Preliminary results suggest recipient government spending on education and health increase almost one for one with education and health sector programme aid, which is entirely on-budget. On the other hand, education and health technical assistance have no effect on recipient sectoral spending, which is in line with expectations that the majority of technical assistance is off-budget. Taken together, these results suggest the fungibility of aid earmarked for education and health is very limited.

*I am grateful to Chris Adam and Paul Collier for useful comments and discussions, and to Ali Abbas, Nicolas Depetris Chauvin, Ibrahim Levent, and Gerd Schwartz for help in obtaining data. I gratefully acknowledge partial funding from St. Catherine's College.

1 Introduction

In 2000 the United Nations General Assembly, then consisting of 189 member countries, adopted the Millennium Declaration, laying the foundations for the Millennium Development Goals (MDGs). The eight Goals 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 *"[make] 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 Official Development Assistance (ODA) and granting debt relief to free resources for social spending (see e.g. United Nations General Assembly, 2006).

Hence, the final MDG recognises the importance of external resources in financing public spending in developing countries. Even when aid finances projects that are carried out without intervention of the recipient government, the latter might adjust its fiscal policy in reaction to the aid inflow. Consequently, the effect of foreign aid on economic growth, poverty, and the targets set out in the MDGs, such as health and education outcomes, depends heavily on the recipient government's fiscal response. One aspect of this fiscal response is the possibility that aid is fungible, which begs the question of how much of the earmarked aid is used for its intended purpose. Is aid given for specific sectors used for spending in those sectors, or is it diverted to other spending programmes (or used to cut taxes or reduce deficits)? In other words: is aid fungible, and if so, to what extent? That is the question this paper endeavours to answer, focussing on the education and health sectors.

Examining the extent of fungibility provides a useful piece of information in gauging the expected effectiveness of aid to reach the MDGs and foster development more generally. As such, studying the fiscal impact of aid, of which fungibility is a part, is at the very least a useful complement to the reduced form relationships between aid and growth that dominate the current academic debate. In addition, fungibility has obvious yet important policy implications for a donor's design of modalities through which aid is delivered. Donors seek to influence fiscal policy choices, and spending choices in particular, by earmarking funds for specific sectors and monitoring compliance. Such earmarking and monitoring is more costly than providing general budget assistance, yet if aid is fungible – as many donors and development scholars seem to believe (see e.g. World Bank, 1998) – both may yield exactly the same outcome. Similarly, if aid is fungible, projects financed and/or carried out by donors outside the realm of the public sector simply release resources recipient governments can use for alternative purposes. In such cases, the success of the project may be a poor guide to the overall developmental impact of aid.

It has long been recognised that aid is potentially fungible. Devarajan, Rajkumar, and Swaroop (1999, p. 1) quote Paul Rosenstein-Rodan (then the Deputy Director of the World Bank's Economics Department) as far back as 1947 as saying: *"when the World Bank thinks it is financing an electric power station, it is really financing a brothel."*[1] Especially in the last two decades scholars have attempted to estimate the extent of fungibility, but this research is not without problems. Most studies lack comprehensive data on the purpose for which aid is given, which is crucial to obtain reliable estimates of fungibility. Drawing on the OECD's Creditor Reporting System, which contains information on the sectoral allocation of aid, this paper constructs measures for the amount of aid disbursements earmarked to the education and health sectors. These sectoral aid variables are linked to recipient public spending in the same sectors in order to provide a more accurate assessment of aid fungibility. To some extent the data also allow to distinguish between on- and off-budget aid flows. This matters, because a failure to recognise that not all aid passes through the recipient's budget leads one to overestimate the extent of fungibility.

In the next section, I define fungibility and provide a brief theoretical motivation for this paper. Next, I review the two main strands in the literature, namely fiscal response models and fungibility studies. Section 4 presents the model to be estimated and discusses the data. This section also contains a simple analytical framework that highlights the importance of distinguishing between aid that goes through the budget and aid that does not. I argue a failure to make this distinction generates an upward bias in the estimated extent of fungibility. Empirical results are discussed in section 5, before concluding in section 6.

[1] Some ten years later, in 1958, Milton Friedman discusses fungibility in his critique of foreign aid published in the Yale Review (reprinted as Friedman, 1964). Because of fungibility, he argues, donors cannot prevent aid from financing economically wasteful projects (monument-building) that fail to support self-sustaining growth in the living standards of the masses.

2 Defining fungibility

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 pure revenue or income augmenting resources that can be spent in the way the recipient chooses (Khilji and Zampelli, 1994). So, even when earmarked aid is fully fungible a small part of it may still be spent on the targeted sector, through a pure income effect. We expect, however, that the marginal allocation of fungible resources towards public education and health expenditure is small, so for practical purposes the difference between both definitions should be small.

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, or to cut taxes and reduce deficits), which we could label general fungibility (McGillivray and Morrissey, 2004). In both cases, the issue is best illustrated graphically.[2]

In figure 1 I assume a recipient government allocates resources between two spending types, say, health and roads. Given the initial budget constraint AB, government utility is maximised at point 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. For simplicity, we assume the additional aid has no effect on the relative price, so the slope of the budget constraint remains unaltered. Left to its own devices, the government now chooses the optimal mix of the two spending types at point F. Earmarked health aid is treated no differently than revenue from other sources, and is thus fully fungible. Hence, fungibility arises as the natural response of a rational government to the inflow of earmarked aid. However, if the donor can force the recipient to spend all of the earmarked health aid within the health sector, aid is non-fungible and the resulting spending allocation is found in point G. This allocation is preferred by the donor, but is suboptimal from the point of view of the recipient government. Such an outcome may be the result of conditions imposed by the donor combined with effective monitoring of aid flows, as discussed in more detail below. If monitoring and conditionality are only partly

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

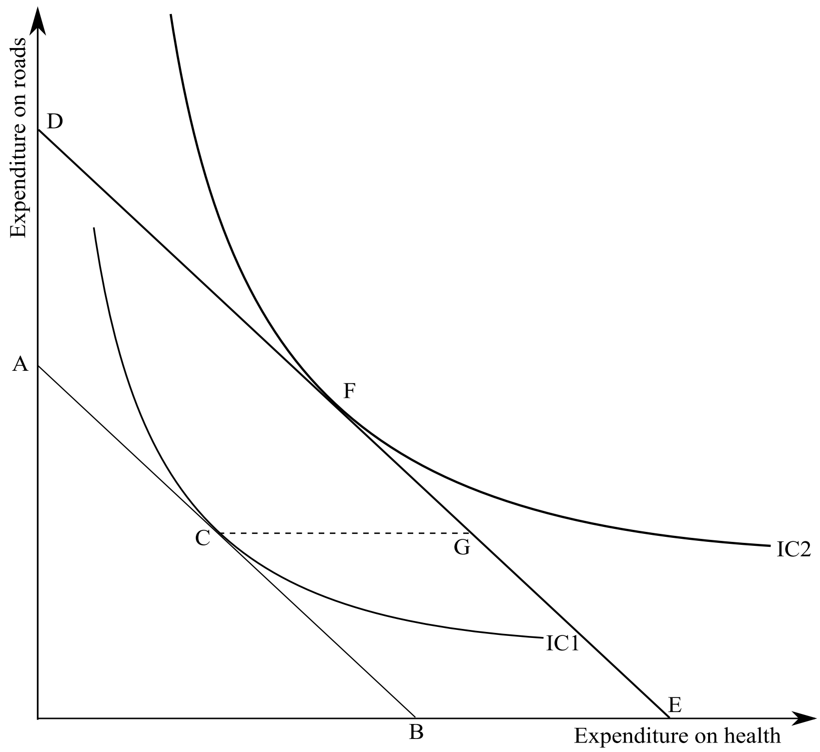

Figure 1: Graphical illustration of aid fungibility

successful, aid is partially fungible, and the resulting spending choice lies between F and G.

This analysis is incomplete in two ways. Firstly, fungibility only arises when aid is diverted from one spending type to another. In reality, governments may also choose to cut taxes or reduce the deficit (increase the surplus) as a response to inflows of aid. As before, such responses imply only part of the earmarked aid feeds through to higher spending in the sector, implying aid is fungible (graphically, the budget constraint DE is pushed back to the left). Secondly, we have dealt only with aid that passes through the government budget. Donors also finance projects in which the recipient government does not intervene. Such aid projects do not show up in the recipient's budget, but may still provoke a fiscal response. For instance, if donors build hospitals on a large scale, the recipient government may cut back its own health spending to compensate. This means we need a broader definition of fungibility: sectoral aid is fungible if total spending in the sector (the sum of government spending and off-budget aid) increases by less than the total amount of earmarked sectoral aid (both on- and off-budget). Figure 1 can be re-interpreted in this light. Simply let the indifference curves reflect government preferences over total spending (government spending and off-budget aid), and let the budget constraint capture all available resources (domestic resources and aid, both on- and off-budget).

2.1 Why aid might not be fungible

Some reasons as to why aid might not be fungible follow immediately from the above analysis. Fungibility crucially depends on a discrepancy in recipient and donor preferences. If preferences are aligned, recipient and donor agree on the sectoral spending allocation at the margin and, as a result, aid is not fungible.

Moreover, the situation described above is essentially that of a one shot game, in which the recipient's optimal strategy is to allocate earmarked funds across sectors according to its own preferences. In a repeated game with a sufficiently long time horizon the outcome may be very different, if the donor is able to monitor the use of aid flows and condition future aid disbursements on the extent of fungibility. Diverting aid yields a contemporaneous increase in utility, which the recipient needs to balance with a potential loss of future utility if donors cut off funds. Such a long term horizon may be relevant if government preferences coincide with those of society as a whole, or if policy makers expect to be in power for a long period of time. If donors can credibly threaten to reduce future aid if fungibility is too high, the extent of fungibility may be seriously curtailed.

One example where monitoring is likely to be effective in reducing fungibility can be illustrated graphically. If donors have information on the total amount of earmarked aid as well as government spending in the sector, and if they can credibly condition future aid flows on the use of current aid, then sectoral spending needs to be at least as high as sectoral aid. This situation produces a kink in the new budget constraint (AHE), as shown in figure 2. This kink reflects the idea that a recipient government can never spend less on a sector than the amount of aid earmarked for that sector (which is CG = AH). Otherwise donors would quickly learn aid is not used for its intended purpose, perhaps prompting them to reduce future aid flows. Because the recipient government knows this, it might prefer to limit the extent of fungibility in order to avoid a reduction in future aid disbursements, in which case the actual budget constraint facing the recipient becomes AHE (instead of DE before).[3] When the amount of health aid exceeds the government's preferred

[3] If earmarked health aid is provided by multiple donors and donors are unaware of the total amount of health aid, the budget constraint is still kinked but AH needs to be interpreted as the largest

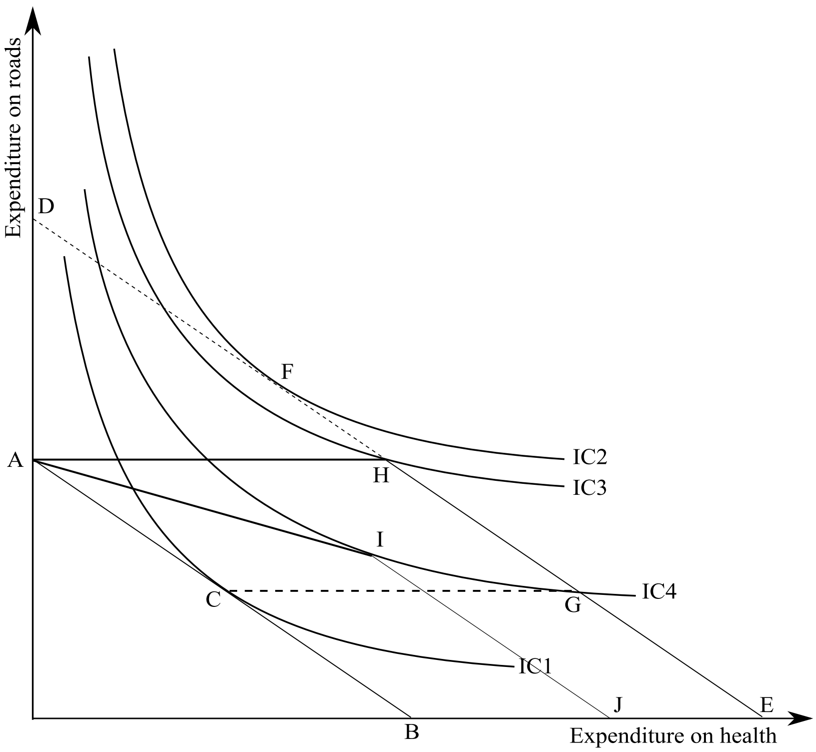

Figure 2: Partial fungibility

amount of health spending given available resources, as is the case in figure 2, fungibility is curbed. Graphically, the kink in the budget constraint (point H) lies to the lower right of point F, which is the recipient's preferred allocation if it did not have to worry about donor conditionality or if monitoring was ineffective (so all resources could be treated as fully fungible). Given budget constraint AHE, the recipient can do no better than the allocation in point H, and health aid is only partially fungible.

Donors may also try to influence the extent of fungibility through the choice of aid modality. Delivering aid as a matching grant, which subsidises the purchase of a good up to a certain threshold, may reduce fungibility. Under a matching grant, the budget constraint becomes AIJ and the government chooses point I as the optimal feasible spending allocation. If the same amount of aid is given as a block grant, entailing a parallel outward shift of the budget constraint to point J

amount of earmarked health aid received from any donor, which sets a lower limit for recipient health spending. If the government spends less than this threshold, health spending falls short of the amount of earmarked health aid provided by the most generous donor (in the health sector), and the latter immediately discovers not all health aid is used for health spending.

(not indicated on the figure), less is spent on health and the extent of fungibility is greater.

3 Literature overview

The previous section suggests theory is ambiguous about the extent of fungibility. Much depends on the ability of a donor to enforce its preferred allocation of aid funds through a combination of earmarking, choice of modality, monitoring, and effective conditionality. Consequently, the extent to which foreign aid is fungible is ultimately an empirical issue.

This section reviews the two main strands of the literature: fiscal response models and fungibility studies.[4] Fiscal response models suffer from a range of problems, raising doubts about the validity of their empirical results. Alternative theoretical models proposed by fungibility studies to guide the interpretation of the empirical results are also based on quite stringent assumptions. This reflects a broader problem in modelling fungibility: because fungibility arises as the natural response of a rational government to the inflow of earmarked aid, some kind of assumption is necessary to "force" the possibility of non-fungibility on the theoretical model. As a result, interpreting empirical results on the basis of these models is potentially misleading. As an alternative guide to interpret regression results I describe a simple analytical framework in the next section when I present the model to be estimated. This mainly serves to pinpoint more precisely a problem in the interpretation of existing estimates of fungibility using donor-based aid data: because not all reported aid passes through the budget of the recipient government, fungibility is overestimated. A potentially even more serious problem is that the bulk of studies lack comprehensive information on the purpose for which aid is given. This paper is the first to use comprehensive cross-country data on the sectoral disaggregation of aid grants and loans to examine fungibility. To a certain degree, this data also enables me to shed light on how the presence of off-budget aid affects estimated fungibility.

3.1 Fiscal response models

Fiscal response models (FRMs) are structural theoretical models that aim to describe a government's fiscal behaviour. Heller (1975) first developed such a model

[4] This classification is borrowed from McGillivray and Morrissey (2004).

more than thirty years ago to investigate the fiscal response to aid. However, it was not until the 1990s that others picked up on his work and FRMs gradually evolved. The current state of the art, as applied in most recent studies, is the model described in Franco-Rodriguez, Morrissey, and McGillivray (1998) and Franco-Rodriguez (2000). I briefly describe this model and highlight where it improves upon earlier FRMs, before pointing out the remaining shortcomings of the model, both in its theoretical foundations and its empirical application.

3.1.1 Theoretical framework

FRMs assume governments make budgetary choices in a rational manner: they set annual targets for fiscal policy variables and try to attain these targets in an attempt to maximise utility under certain constraints. In Franco-Rodriguez, Morrissey, and McGillivray (1998) the following quadratic loss function is maximised:

$$\begin{aligned} U = \alpha_0 - \frac{\alpha_1}{2}(I_g - I_g^*)^2 - \frac{\alpha_2}{2}(G - G^*)^2 - \frac{\alpha_3}{2}(T - T^*)^2 \\ - \frac{\alpha_4}{2}(A - A^*)^2 - \frac{\alpha_5}{2}(B - B^*)^2 \end{aligned} \tag{3.1}$$

where I_g is public capital expenditure (investment), G is government consumption, T is total revenue, B is borrowing from domestic sources and A is aid. The asterisks denote exogenous targets. $\alpha_i > 0$ for $i = 1, \ldots, 5$, so utility falls as targets are missed by a wider margin. The marginal utility gained from moving towards the target decreases as the deviation from the target becomes smaller, implying a diminishing marginal utility of moving closer to the target. Maximum utility α_0 is attained when all targets are met.

This model departs from its predecessors in that Franco-Rodriguez, Morrissey, and McGillivray (1998) endogenise aid by including it in the utility function as a choice variable. They argue that, while donors typically determine the amount of aid committed, it is ultimately the recipient who decides how much of the commitment is disbursed in each year. So, in (3.1) A reflects aid disbursements and A^* is the amount of aid committed. This assumption is questionable, and no real evidence is presented to support it, apart from an observation that aid commitments and disbursements often differ markedly. Moreover, even if governments exert a great deal of influence on aid disbursements it is unlikely A^* coincides with aid commitments. Franco-Rodriguez, Morrissey, and McGillivray (1998) argue deviations from commitments in either direction cause disutility, as $A < A^*$ reflects limited absorptive capacity, while $A > A^*$ reflects disbursements of previously undisbursed commitments (limited absorptive capacity) or emergency aid (which proxies for an adverse shock). This may be true, but many other considerations could lead the target for aid to diverge from the amount of aid committed.

Utility is maximised subject to the following constraints:

$$G \leq \rho_1 T + \rho_2 A + \rho_3 B \tag{3.2}$$

$$I_g + G = T + A + B \tag{3.3}$$

where ρ_1, ρ_2 and ρ_3 are the proportions of total revenue, aid and borrowing directed towards consumption spending (I further refer to the ρs as fungibility parameters). At first blush ρ_2 thus reflects to what extent aid is fungible (though see section 3.1.2 below). Equation (3.3) is the government budget constraint. Equation (3.2) captures an external constraint on fiscal policy choices. The underlying idea is that domestic interests and donors impose the ρs on policymakers and, as a consequence, targets may not be met even if (3.3) is satisfied.[5] Franco-Rodriguez, Morrissey, and McGillivray (1998) set up a Lagrangian for this problem and derive the first order conditions, assuming (3.2) is binding. If (3.2) does not hold with equality, utility is maximised subject only to the general budget constraint, yielding an internal solution (Franco-Rodriguez, 2000). It is then optimal to set the fiscal policy variables equal to their target, unless the budget constraint is not satisfied at the targets, which may occur if targets are set independently from each other.

From the first order conditions for the corner solution a set of structural equations can be found, expressing each fiscal policy variable in terms of exogenous targets and endogenous variables. Because the right hand sides of the structural equations still contain endogenous variables, these equations cannot be used to study the full effect of aid that works through all elements of the budget. The coefficients in the structural equations are non-linear combinations of the ρs and αs. Further solving this system of structural equations yields reduced form equations of the following form (see Mavrotas and Ouattara, 2006, for the correct derivation):

[5] McGillivray and Ouattara (2005) focus on a variant of this model with debt service on the left hand side of (3.2). Hence, they study fungibility between total government expenditure and debt service rather than between public investment and consumption.

$$I_g = \pi_1 I_g^* + \pi_2 G^* + \pi_3 T^* + \pi_4 A^* \quad (3.4)$$

$$G = \pi_5 I_g^* + \pi_6 G^* + \pi_7 T^* + \pi_8 A^* \quad (3.5)$$

$$T = \pi_9 I_g^* + \pi_{10} G^* + \pi_{11} T^* + \pi_{12} A^* \quad (3.6)$$

$$A = \pi_{13} I_g^* + \pi_{14} G^* + \pi_{15} T^* + \pi_{16} A^* \quad (3.7)$$

$$B = \pi_{17} I_g^* + \pi_{18} G^* + \pi_{19} T^* + \pi_{20} A^* \quad (3.8)$$

where B^* is assumed zero and π_1 to π_{20} are again non-linear functions of the underlying ρs and αs. These reduced form equations do capture the full effect of aid on fiscal policy variables. Studies that do not endogenise aid would have A instead of A^* on the right hand side. In addition, such studies typically employ more disaggregated expenditure and aid categories, so that, for instance, fungibility is allowed to differ between grants and loans. Endogenising aid increases the number of coefficients to be estimated, so to keep the model tractable Franco-Rodriguez, Morrissey, and McGillivray (1998) do not further disaggregate aid and spending, acknowledging this may lead to an aggregation bias if the relative weights of the disaggregated variables in the utility function differ strongly. In a sense, not endogenising aid is also intuitively more attractive because it allows to look at the effects of aid disbursements, which seems more relevant than the effects of aid commitments.

Given information on the exogenous targets, most studies use non-linear Three Stages Least Squares (3SLS) to estimate the system of structural equations (structural equations are not reported here for sake of brevity). The latter technique is appropriate as the structural equations form a simultaneous system with cross-equation restrictions on the parameters. By substituting the estimated parameters in (3.4)-(3.8), indirect estimates of the reduced form effects are obtained.

This model solves some problems encountered in earlier FRMs (Heller, 1975; Gang and Khan, 1991; Khan and Hoshino, 1992) that specify a linear-quadratic loss function instead of (3.1), which – in the case described here – would take on the following form:

$$
\begin{aligned}
U = {} & \alpha_0 + \alpha_1(I_g - I_g^*) - \frac{\alpha_2}{2}(I_g - I_g^*)^2 + \alpha_3(G - G^*) - \frac{\alpha_4}{2}(G - G^*)^2 \\
& - \alpha_5(T - T^*) - \frac{\alpha_6}{2}(T - T^*)^2 + \alpha_7(A - A^*) - \frac{\alpha_8}{2}(A - A^*)^2 \\
& - \alpha_9(B - B^*) - \frac{\alpha_{10}}{2}(B - B^*)^2
\end{aligned} \tag{3.9}
$$

In this specification utility is no longer maximised when the targets are attained (Binh and McGillivray, 1993). The government is better off to overshoot investment, consumption and aid, and to undershoot revenue and borrowing. Therefore, targets can no longer truly be interpreted as targets, and the structural equations derived are inconsistent with maximising behaviour. As an alternative, Binh and McGillivray (1993) propose a quadratic loss function, noting a potentially restrictive characteristic of this utility specification is its symmetric nature: overshooting and undershooting impose the same utility loss, which may not be realistic. More generally, it is unclear whether the loss function adequately represents how governments actually behave (McGillivray and Morrissey, 2004).

More importantly, Franco-Rodriguez, Morrissey, and McGillivray (1998) invoke a different set of constraints. Some other studies substitute (3.3) and (3.2) for constraints of the form (Heller, 1975; Gang and Khan, 1991; Khan and Hoshino, 1992; Mavrotas, 2002):

$$I_g = (1 - \rho_1)T + (1 - \rho_2)A + B \tag{3.10}$$

$$G = \rho_1 T + \rho_2 A \tag{3.11}$$

which decompose the overall budget constraint. One problem with these constraints is they impose borrowing cannot finance consumption spending. This could easily be addressed by replacing B in (3.10) by $(1 - \rho_3 B)$ and adding $\rho_3 B$ to (3.11). A more serious problem is this set-up over-constrains the model, meaning α_0 may not be achieved even if both constraints are satisfied (White, 1994). This arises because, even if total resources suffice to finance total spending, the ρs constrain how resources can be allocated to specific expenditure targets. These ρs are not necessarily at the value required to achieve the optimal allocation of resources because they are predetermined, rather than being determined as the outcome of a utility maximisation problem. Though the model is no longer over-constrained with constraints (3.2) and (3.3), the core of the problem – that the ρs are predetermined – is not resolved. Replacing both constraints by a single budget constraint also solves the over-constraining problem (White, 1994), but implicitly assumes all aid is fully fungible, because it does not incorporate any restrictions on how aid is allocated across different expenditure categories. Hence, it cannot be employed to study fungibility.

Constraints (3.2) and (3.3) are not without drawbacks either, though. The inequality constraint is essentially ad hoc. Few compelling reasons are given as to why we would expect external pressure from politicians, interest groups, donors, elements of the bureaucracy,... to be exercised in such a way that it limits the amount of revenue, aid and borrowing that can be allocated towards public consumption. In fact, bureaucracies and voters might exercise exactly the opposite type of pressure and push for higher public consumption (for instance, in the form of higher wages and social transfers). Similarly, in recent years donors have begun to realise a lack of public funds for operations and maintenance may seriously reduce the returns to investment projects. Moreover, government spending on health and education is often classified as consumption. As such, donors might believe government consumption (or at least, certain aspects of it) in the recipient country is too low, which could change the direction of the inequality.

3.1.2 Empirical application

More problems arise in the empirical application of the above described FRM.

Firstly, non-linear least squares estimation is very demanding of the data and the results it yields are not very robust (McGillivray and Morrissey, 2004).

Secondly, ρ_2, the parameter of interest, reflects the *maximum* rather than the exact amount of fungibility. This is because the rather implicit assumption that all aid is intended for public investment may not always hold. In reality, especially in more recent years, donors may intend aid flows to partly finance government consumption, in which case ρ_2 is expected ex ante to lie above zero even if aid is not fungible (Franco-Rodriguez, Morrissey, and McGillivray, 1998). This is especially the case because government spending on education and health is generally not recorded as public investment. Moreover, if (3.2) does not hold with equality, most likely less than ρ_2 is diverted towards consumption (Franco-Rodriguez, 2000). So, again, ρ_2 would exaggerate the actual extent of fungibility. This reflects a more general problem that one needs to assume (3.2) is binding, for only in that case can structural and reduced form equations be derived whose coefficients link back to the fungibility parameter. Franco-Rodriguez (2000) argues no econometric technique is available to model an endogenous switching between the equation systems of the interior and the corner solution.

Thirdly, no data is available for the exogenous targets. Therefore, targets are approximated as the fitted value from a regression of the actual variable on a small number of presumably exogenous regressors. For instance, Franco-Rodriguez, Morrissey, and McGillivray (1998) estimate I_g^* as the fitted value from a cointegrated regression of I_g on lagged private investment, lagged GDP, and lagged public sector borrowing requirement. Alternatively, the specification for this first stage regression can be substituted in for the targets in the equation system. White (1994) notes this is highly problematic. If the fit of the first stage is very good, and the R^2 approaches 1, I_g is very close to I_g^* and regressing the former on the latter in the structural equation gives a coefficient close to 1, whereas the other coefficients will be insignificant. If the fit of the first stage is bad the fitted value cannot reasonably be interpreted as a target variable: either the wrong independent variables are included or the actual value is far away from its target, in which case the estimated coefficients are not the ones that are relevant in determining the target. Moreover, for some variables targets are linked to past realisations of the variable. For instance, Gang and Khan (1991) let G_c^*, the target for spending on the maintenance of political and bureaucratic organisations, depend on lagged G_c. Hence, any contemporaneous effect of aid on G_c changes G_c^* in the subsequent period, which in turn affects all endogenous variables in that period, and further feeds through to later years. This introduces an implicit dynamic element in the model, which is suppressed in estimation (White, 1994).

Fourthly, many studies omit reduced form equations, focusing on the effects found in the structural equations and/or the estimated ρs (the fungibility parameters). As already discussed, because the structural equations still contain endogenous variables on the right hand side, coefficients in these equations only capture partial effects. The most relevant results lie in the reduced form equations, as only these equations capture the full effect of aid feeding through all elements of the budget. White (1994) and McGillivray (1994) discuss how these reduced form results may differ substantially from the conclusions reached on the basis of fungibility parameters or partial effects in the structural equations.

Fifthly, to the extent that FRMs rely on donor-based aid data an unknown fraction of aid is off-budget (that is, it does not pass through the recipient's budget), which may affect results.

Lastly, and perhaps a reflection of the general problems of FRMs, in many studies the estimated ρs lie outside the $[0, 1]$ bound. Heller (1975) interprets a negative fungibility parameter for total loans (the sum of official and private loans) as indicating that a dollar of total loans pulls non-loan resources from the recurrent budget. Similarly, Gang and Khan (1991, p. 363) conclude that *"the proportion of tax revenues that remain in the current budget is 108%"*, arguing this reflects funds being pulled out of development projects when taxes are raised. Theoretically, however, all ρs should lie between 0 and 1, as one cannot spend a negative amount of aid (or taxes) on a certain expenditure category, nor can one spend more than the amount of aid received or taxes raised. McGillivray and Ouattara (2005) restrict the fungibility parameters and other parameters in the structural equations to lie within the bounds postulated by theory. However, it is doubtful whether this is the right solution. Rather, the frequent violation of parameters to lie within the bounds postulated by theory suggests fiscal response models are not supported by the data. Hence, FRMs as currently applied by scholars are perhaps not an adequate way to study the fiscal response to aid.[6]

Because of the many problems plaguing these studies, at this stage I do not discuss their empirical findings. I will, however, briefly contrast the general pattern of results found in FRMs with those found in fungibility studies after I have discussed the latter (see section 3.2.2).

3.2 Fungibility studies

Fungibility studies estimate single equations or systems of equations without relying on a structural theoretical model to provide an all-encompassing description of budgetary choices. Thus, their main unifying characteristic is that they are not fiscal response models. Some fungibility studies develop alternative theoretical models to guide the interpretation of empirical results and to allow fungibility parameters to be recovered from estimated coefficients. Of these, I briefly describe the model in Feyzioglu, Swaroop, and Zhu (1998). This model is based on strong assumptions and may thus cloud, rather than enhance, understanding of the empirical results, possibly yielding misleading fungibility parameters. Hence, despite the fact that the empirical set-up employed in this paper is close to that of the fungibility studies discussed in this section, I refrain from interpreting coefficients in light

[6]Recently, Osei, Morrissey, and Lloyd (2005) have estimated the fiscal response to aid in Ghana using a cointegrated VAR model. This is a promising new approach, which allows to move away from some of the stringent assumptions inherent in fiscal response model and circumvents the need to estimate fiscal policy targets, while still being able to trace the effect of aggregate aid through the budget of the recipient country.

of existing theoretical models. I discuss the empirical results found in fungibility studies in quite some detail, mainly because these studies suffer from a range of problems which are useful to bear in mind when interpreting results. Moreover, highlighting these problems is helpful to guide the empirical design of this paper, and to clarify its contribution in the literature.

3.2.1 An alternative model of fungibility (Feyzioglu, Swaroop, and Zhu, 1998)

A recipient government purchases S public goods to provide to its citizens.[7] Aid is earmarked by purpose towards the purchase of $K \leq S$ goods, a_k $(k = 1, \ldots, K)$ being the amount of aid intended for each good k. A portion $0 \leq \phi_k \leq 1$ of the earmarked aid for good k is fungible. This fungible portion of aid can be treated as a lump-sum addition to government revenue net of aid, denoted R. The quantity of good s paid for with the fungible portion of aid and revenue from other sources (R) is denoted as g_s $(s = 1, \ldots, S)$. g_k^{NF} is the quantity of good k that must be purchased from the non-fungible part of aid. The representative agent's utility function depends on the S public goods and a single private good c_p:

$$\begin{aligned} W = U[c_p, g_1, g_1^{NF}, \ldots, g_K, g_K^{NF}, g_{K+1}, \ldots, g_S] \\ \text{with } g_k^{NF} = \frac{(1-\phi_k)a_k}{p_k} \text{ for } k = 1, \ldots, K \end{aligned} \tag{3.12}$$

where p_s is the price of good s. To allow for an analytical solution utility is assumed to be of the Stone-Geary form. The government maximises W subject to the following budget constraint, in which only fungible resources appear:

$$p_1 g_1 + p_2 g_2 + \ldots + p_s g_s = R + \sum_{k=1}^{K} \phi_k a_k \tag{3.13}$$

If an interior solution exists this maximisation problem yields a system of linear expenditure equations (one for each good $s = 1, \ldots, S$). These can further be manipulated to yield equations in which spending on each good depends on the aid earmarked for that good, the fungible part of aid earmarked for other goods, and total government spending net of foreign aid (see Feyzioglu, Swaroop, and

[7]The notion of purchasing public goods on the market, which is the starting point of the model in Feyzioglu, Swaroop, and Zhu (1998), seems somewhat odd. However, the model is equally valid if we assume public goods are produced by the government, and the prices introduced later reflect production costs.

Zhu, 1998, for details). Empirical versions of these equations can be estimated, and fungibility parameters for each sector recovered from the estimated coefficients.

A crucial assumption of the model is that public goods purchased from the non-fungible component of aid are a separable argument in the utility function. This means the non-fungible part of aid does not affect the choice of public goods purchased from the fungible budget resources ($g_1, \ldots, g_s$). The latter are only affected through the fungible portion of aid. Suppose a government receives a large amount of non-fungible health aid. Examples could be donor-financed health projects that bypass the recipient government, or other forms of off-budget health aid. A natural response by the government is to limit spending on health from fungible resources in order to compensate for the fact that donors are building hospitals, supplying syringes, and paying doctors and nurses. Otherwise total resources devoted to health (the sum of public health spending and off-budget health aid) would exceed the amount deemed optimal by the recipient government, which contradicts the simple analysis in section 2. Yet it is precisely this rational adaptation by the government of its spending from fungible resources that is ruled out a priori by the assumption Feyzioglu, Swaroop, and Zhu (1998) make. If this assumption is not made, the resulting spending choices would imply full fungibility as in figure 1. A second problem with the model is that, as in FRMs, fungibility parameters are not determined within the model, but simply imposed as exogenous parameters.

The drawbacks of the theoretical model become obvious when looking at some of Feyzioglu, Swaroop, and Zhu's (1998) results. Several of the fungibility parameters, derived from the estimated coefficients, lie outside the $[0, 1]$ bounds postulated by the theoretical model, often by a wide margin (the smallest fungibility parameter is -3.96, the largest 45.44).

Feyzioglu, Swaroop, and Zhu's (1998) model is reminiscent of an earlier model developed by McGuire (1978). McGuire (1978) decomposes the effect of a targeted grant into an income and a price effect. The income effect arises because part of the grant is fungible, adding to the fungible resources available in the economy. The price effect comes from the non-fungible component: this part is channelled directly to the targeted sector and reduces the effective price paid by the recipient, as a greater quantity of the targeted good can be bought with the same amount of expenditure from fungible resources devoted to the good. Hence, the idea is that the observed post-grant allocation is consistent with different pairs of effective post-grant budget constraints and indifference maps. Each post-grant budget constraint encapsulates a different combination of income and price effect, and, hence, a different extent of fungibility (also see Zampelli, 1986). So, the aim is to find which pair best fits the data.

Formally, the average citizen's (Stone-Geary) utility, which depends on per capita quantities of various public and non-public goods, is maximised subject to a budget constraint which sets total expenditure (quantities multiplied by grant-induced effective prices, summed over all goods) equal to fungible resources available in the economy (sum of own fungible resources and the fungible components of targeted aid). From this a linear expenditure system is derived and taken to the data.

Apart from the restrictions imposed by having to choose a specific form for utility, McGuire's (1978) model assumes the post-grant observed allocation is at an optimal point, i.e. a point where an indifference curve touches a budget constraint. By confronting the model with the data under this assumption, the pair of post-grant budget constraints and indifference maps which best fits the data is identified. This requires that the government agency under investigation behaves as if it is maximising utility of the representative agent, which may not be the case. Perhaps more importantly in this context, the main effect of aid conditionality may be to force the recipient government to choose a suboptimal outcome (see figure 1 on page 4), which would again violate this assumption.

McGuire (1978) applies this model to examine the effect of US intergovernmental education expenditure (this is the sum of federal and state aid) on state-local education expenditure, and finds about 70% of education grants are fungible. Employing a comparable model, Zampelli (1986) observes similar high fungibility parameters for US federal grants for social services and urban support services directed to 18 large cities. However, for the third category (all other direct general government aid) he concludes aid is not fungible. Moreover, a separate fungibility parameter for state aid is found to be insignificantly different from zero, but this may be influenced by the untested restriction that state aid fungibility is the same for all three categories. Results of studies that apply this model to foreign aid are discussed in the next section.

3.2.2 Empirical results

The earliest fungibility studies are applications of McGuire's (1978) model. McGuire (1982; 1987) applies the model to time series data for US aid given to Israel (1960-1979), adding supply side equations to explain US assistance and Arab defense expenditure (this is the sum of real defense expenditure of Israel's neighbour countries). The latter enters the utility function through the subsistence quantities, the idea being that the magnitude of the Arab threat impacts on the optimal allocation of Israeli resources. McGuire (1982) finds US economic assistance to Israel is fully fungible (90-100%) but only a small part of military assistance is fungible (4-18%). McGuire (1987) in addition distinguishes private consumption goods from investment goods, now finding that both economic and military assistance are highly fungible. Khilji and Zampelli (1991) reach the same conclusion for defense and non-defense US aid to Pakistan (1960-1986). In both McGuire (1987) and Khilji and Zampelli (1991) fungible resources are predominantly used for tax relief. Khilji and Zampelli (1994) employ a panel of 8 countries (1972-1987) that receive aid from the US, and find no evidence to suggest fungibility differs across countries. Military assistance is fully fungible, while the fungibility parameter for non-military assistance is significantly below one though still high (0.65).

Much like FRMs, estimating McGuire's (1978) model is demanding of the data and a large number of parameters are typically estimated on small samples. As such, authors often impose prior restrictions to enable identification of the remaining parameters. McGuire (1982), for instance, estimates the supply side equations without intercepts. As a result, these models are expected, like FRMs, to be fragile. The change in the estimated fungibility for military assistance between McGuire (1982) and McGuire (1987) is perhaps testimony to this. As in FRMs and in Feyzioglu, Swaroop, and Zhu (1998), estimated fungibility parameters are not always within the [0-1]-interval predicted by theory.

A second issue is the assumption that all targeted aid (e.g. military assistance) is recorded in the receiver's budget as targeted (military) expenditure. In the expenditure equations the dependent variables are *local* expenditure on the different goods, calculated as total sectoral expenditure minus aid targeted to that sector. So, for instance, local military expenditure is calculated as total military spending minus US military assistance. As a result, to the extent military assistance is not fully recorded as military expenditure, local expenditure is underestimated and fungibility is over-estimated.

Other studies also focus on the military-non-military divide, but in a reduced form framework. Stein, Ishimatsu, and Stoll (1985) find US military aid programmed in the previous year substitutes for a recipient's own military expenditure (expressed as a share of total central government expenditure) in the current year. Zahariadis, Travis, and Diehl (1990) find the effect of US economic aid on a recipient's military expenditure depends on the type of program. While Economic

Support Funds (ESF) increase military expenditure, for Development Assistance (DA) and other aid the coefficient is negative. According to the authors, ESF is intended to free domestic resources that can be directed towards military expenditure, so fungibility is almost inherent in the program, whereas for the other aid types security is much less of a concern, implying fungibility is expected to be lower.

The studies reviewed so far focus narrowly on the military-non-military divide, and on aid received from only one donor (the US). Not further disaggregating non-military aid might mask fungibility between the disaggregated sectors. More generally, it is difficult to assess to what extent results generalise to non-military sectoral aid categories (health, education, infrastructure,...). The strategic relationships between the US and the recipients under study, for instance, might be driving results. As discussed by Zahariadis, Travis, and Diehl (1990), part of US non-military aid is intended to free up resources in the recipient country that can be used for military expenditure, so in this case the US intends aid to be fungible. Finally, these studies rely on donor-based aid data, part of which may be off-budget.

In order to provide a more general assessment of fungibility, a few studies make use of an unpublished World Bank database, the Debtor Reporting System (DRS), that enables to split up concessionary loans according to the sector the loans are intended for. Categorical fungibility is then assessed by comparing the marginal effect of sectoral loans on recipient public spending in the same sector with the marginal effect of domestic resources. If the marginal effect of sectoral loans is equal to or smaller than the marginal effect of government expenditure net of aid, loans are fully fungible. Feyzioglu, Swaroop, and Zhu (1998) find that, for four out of six sectors, the effect of sectoral aid is insignificant. Only for transport and communication the coefficient is close to 1 (no fungibility), while energy loans are partly fungible (coefficient 0.36, which exceeds the marginal effect of government expenditure net of aid).[8] Devarajan, Rajkumar, and Swaroop (1999) reach similar conclusions for 18 Sub-Saharan African countries (1971-1995). Sectoral loans have a positive impact on spending in two sectors: energy, and transport and communication. In both cases the effect is a lot smaller than 1 but larger than the effect

[8]As argued in section 3.2.1, calculation of the fungibility parameters from the estimated coefficients builds on a model that is based on a strong assumption. Hence, I focus on the marginal effects of sectoral aid on sectoral spending to assess whether aid is fungible. When the sectoral equations are jointly estimated Feyzioglu, Swaroop, and Zhu (1998) only report the fungibility parameters. The main result in that case is again that only transport and communication aid is not fungible, while the fungibility parameters in the other sectors are very sensitive to the choice of government expenditure data and often take on extreme values (see section 3.2.1).

of domestic revenue, indicating partial fungibility. Concessionary loans have no clear effect in the other sectors. Both papers find some evidence to suggest the effect of aggregate aid on total government spending is strong (close to 1), and that most of this is accounted for by increases in public consumption. Using time series data on India (1970-1995) Swaroop, Jha, and Rajkumar (2000) similarly find aid increases non-development expenditure, but has no impact on development expenditure or revenue. A sectoral disaggregation for concessionary loans is again obtained from the same World Bank database. For none of the seven development expenditure categories a significant effect of sectoral loans is found.

The main advantage of using the DRS data is that this way at least some information is available to discern the sector for which aid is intended. In addition, since the data is debtor-based, all reported loans should pass through the recipient's budget. However, an obvious drawback is that these measures of earmarked aid are very incomplete, because grants are omitted. In addition, if the quality of donor-based aid data is superior to that of debtor-based data, the latter could contain more measurement error.

Pack and Pack (1990; 1993) are able to go beyond this, by focusing on single countries for which both sectoral aid and sectoral public spending data are available. Both papers estimate a system of equations for non-development current expenditure, several types of development expenditure and government revenue (excluding foreign aid) with Seemingly Unrelated Regressions (SUR), using time series data for Indonesia (1966-1986) and the Dominican Republic (1968-1986), respectively. Pack and Pack (1990) find that, for all five categories of development expenditure, the marginal effect of sectoral aid on sectoral spending is close to 1, suggesting fungibility is very limited. Pack and Pack (1993), however, find results that are much less encouraging for donors. Only in one sector (health, education, and social services) the marginal effect of categorical aid on categorical spending is close to 1. For some development expenditure categories the effect is negative. Development expenditure and especially total expenditure (which also includes current expenditure, and financial and indirect investment) are lower than before the aid inflow (total expenditure is 27 cent lower per dollar of aid). Aid is mainly used for deficit reduction and debt repayment (88 cents per dollar of aid) and also to some extent to finance lower tax revenue.

Results from fungibility studies, far from being conclusive, seem to imply aid is at least partially fungible. Though some studies find a relatively strong positive effect on total spending for some samples, the bulk of this is accounted for by increases in public consumption and it is unlikely this is what donors intended. Moreover, typically aid is categorically fungible, with Pack and Pack (1990) as the only notable exception.

I have not yet discussed the empirical results found in FRMs. Given the many problems these studies face, results should be taken with a grain of salt. Moreover, results but also the exact model used and the parameters or effects on which conclusions are based differ between studies. All this makes it difficult to distill a general pattern of results for FRMs. Remaining cautious, perhaps the main result is that the FRMs as a whole show more mixed results on general fungibility. That is, the evidence that aid is predominantly spent on consumption and not on investment is not compelling in this models, though at the same time no consistent evidence to suggest significant increases in public investment is found (McGillivray and Morrissey, 2000; McGillivray and Morrissey, 2004).

4 Empirical model and data

4.1 Empirical model

I estimate reduced form regressions linking earmarked sectoral aid to sectoral public spending in the recipient country. Hence, the spirit of the paper is close to that of Pack and Pack (1990; 1993), the main difference being I employ cross-country data rather than a single time series. In addition, I estimate single equations for two sectors (education and health), rather then estimating a system of equations of fiscal policy variables. As such, the empirical set-up is also close to that of Feyzioglu, Swaroop, and Zhu (1998) and Devarajan, Rajkumar, and Swaroop (1999). Differences are that I do not derive a fungibility parameter from the estimated coefficients and, more importantly, that I use more complete measures of sectoral aid.

This paper estimates versions of the following model:

$$SSP_{i,t} = \alpha + \beta SAID_{i,t} + \sum_{k=1}^{K} \gamma_k A_{i,t}^k + \sum_{l=1}^{L} \delta_l X_{i,t}^l + \eta_i + \epsilon_{i,t} \tag{4.1}$$

where η_i captures country-specific fixed effects and $\epsilon_{i,t}$ is the transient error, assumed iid. $SSP_{i,t}$ denotes sectoral spending on education or health, while $SAID_{i,t}$ is aid intended for the same sector. β is thus the main parameter of interest. $A_{i,t}^k (k = 1, \ldots, K)$ and $X_{i,t}^l (l = 1, \ldots, L)$ denote other aid measures and control variables, respectively, as described in more detail in section 4.2. Aid and spending variables are expressed as a share of GDP. Scaling by GDP better reflects the opportunity costs of alternative uses of resources than when the variables are expressed in real per capita terms. Hence, this specification is theoretically more appealing since it better captures the trade-offs the government faces. Moreover, since spending as a share of GDP is bounded between zero and one, non-stationarity is less likely to be a problem, reducing the probability that our results are spurious.

I focus on education and health for two reasons. Firstly, education and health play a prominent role in the MDGs. Apart from their importance in eradicating extreme poverty, five other Goals explicitly set targets related to education and health.[9] 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, these are rather clearly defined areas of spending, which should increase the definitional overlap between sectoral aid and sectoral spending.

Several things should be noted about (4.1). First of all, our linking of sectoral aid to sectoral spending constitutes a marked improvement on most previous studies that lack information on the purpose for which aid is given. Feyzioglu, Swaroop, and Zhu (1998), Devarajan, Rajkumar, and Swaroop (1999) and Swaroop, Jha, and Rajkumar (2000) are exceptions, but they are only able to disaggregate concessionary loans according to sector, so the omission of sectoral grants may influence their results. Pack and Pack (1990; 1993) are the only studies that employ a comprehensive sectoral disaggregation of both grants and loans, but these studies suffer from other serious problems. Both focus on a single country time series but fail to examine the time series properties of the data, simply adding a trend to some of the equations. Therefore, results are potentially spurious, especially since nominal per capita data is used. One indication for this is that, in the 1990 study, the R^2 never falls below 0.74, while in Pack and Pack (1993) none of the equations have an adjusted R^2 below 0.98. Perhaps even more seriously, both estimate a system of equations using SUR on a very small sample of 21 and 19 observations, respectively.

As such, this is the first cross-country study that makes use of a comprehensive sectoral disaggregation of grants and concessionary loans to study fungibility.

[9]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 General Assembly, 2006).

Second, I do not control for other fiscal variables on the right hand side of (4.1). Most notably, no measure of domestic resources net of aid is included. Including such a measure would eliminate part of the effect of sectoral aid, which might run through changes in total spending from domestic resources. This could be circumvented by following the procedure in Devarajan, Rajkumar, and Swaroop (1999) and including that part of domestic resources net of aid that is orthogonal to the included aid variables. This strategy makes a correct assessment of the extent of fungibility possible, as it allows to compare the effect of sectoral aid with the effect of domestic resources net of aid (again, see section 3.2.2). However, because data availability for total spending (taken from International Monetary Fund, 2006) is limited a great deal of information would be lost, so I choose to omit domestic resources. This is unlikely to be of great practical significance. Unless there is a substantial break in policy, the marginal effect of domestic resources should be close to its average allocation. This implies that for most sectors, including education and health, the marginal effect should be close to zero. This is indeed what Devarajan, Rajkumar, and Swaroop (1999) find: for health spending the coefficient of domestic resources is 0.04, for education spending it is 0.12 (spending and aid measured in real per capita dollars). Feyzioglu, Swaroop, and Zhu (1998) find even smaller effects.

Third, it is important to make clear the value we expect for β under different scenarios. Because of the cross-country nature of our analysis, aid variables are necessarily obtained from a donor-based database. McGillivray and Morrissey (2000, p. 422) argue that, because a large portion of aid measured by donors does not go through a recipient's public sector account, such aid measures *"... are inappropriate for analysing fungibility"*. However, consistent with the extended definition of fungibility discussed in section 2, the presence of off-budget aid simply calls for a reinterpretation of β.

Table 1: Approximate expected value for β under different scenarios

	aid not fungible	aid fungible
aid off-budget	0	-1
aid on-budget	1	0

Table 1 summarises the expected values for β under different scenarios. When aid is on-budget and not fungible, the effect of sectoral aid on sectoral spending should be at least 1. When aid is on-budget and fungible, β should equal the effect of domestic resources net of aid, and thus – as already discussed – be close to 0. When donor aid is off-budget, aid is non-fungible when the total amount of resources devoted to the sector (the sum of government spending and off-budget aid) increases by – at least – the amount of off-budget aid. Hence, for off-budget aid non-fungibility is consistent with $\frac{\partial SSP}{\partial SAID} = \beta = 0$. Fungibility arises if the total amount of resources spent on the sector exceeds the government's preferred amount. In that case the government tries to compensate for the increase in off-budget aid by reducing spending from domestic resources on the relevant sector. So, fungibility in combination with off-budget aid leads to $\beta < 0$, and full fungibility is associated with $\beta \approxeq -1$. Hence, the presence of off-budget aid would lead to overestimate the extent of fungibility. Keeping this interpretation in mind, we should still be able to say something useful about the extent of fungibility. Moreover, as discussed later, I split up sectoral aid in four different components. For each we expect a different proportion to go through the budget. This allows us to make more precise statements about the extent of fungibility.

4.2 Data

Education and health spending as a share of GDP are IMF staff estimates from the IMF's Fiscal Affairs Department.[10] The latest update for this data, after which collection was ceased, is October 2004, so 2003 is the last year for which data is available (I received the data December 2006). The data is originally taken from IMF country documents, which are then checked by desk economists for each country to verify and reconcile the data (Baqir, 2002). The main advantage over other datasets (International Monetary Fund, 2006; World Bank, 2006a; World Bank, 2006c) is its greater coverage. Also, the level of government to which the data refer differs across countries, but is fixed across time, so these differences in government level are picked up by fixed effects (Baqir, 2002). This data has been used in a variety of publications (Gupta, Clements, and Tiongson, 1998; Gupta, Dicks-Mireaux, Khemani, McDonald, and Verhoeven, 2000; Baqir, 2002; Thomas, 2006; Lora and Olivera, 2006).

Aid data is constructed from the two databases in the OECD's International Development Statistics online: the Development Assistance Committee (DAC) database and the Creditor Reporting System (CRS) database. The former is described in OECD (2000a), the latter in OECD (2002). All aid data is reported in millions of US$.

[10] I am very grateful to Gerd Schwartz for sharing this data, and to Ali Abbas for help in obtaining the data.

The CRS database contains Official Development Assistance[11], aggregate but also disaggregated by purpose (sector). Gross disbursements are available for 1990-2004. This data can be obtained in a recipient-donor-year format, i.e. for each year it shows how much aid is transferred from a given donor to a given recipient. Unfortunately, CRS disbursements reported by some multilateral and bilateral donors are incomplete or even nonexistent for some years.

Hence, I further use total aid disbursements from DAC table 2a, again in a recipient-donor-year format. This data should be complete, but it does not allow to split up aid by purpose. I calculate, for each recipient-donor-year observation, the amount of "missing" aid, as the difference between DAC2a and CRS disbursements. The aim is to allocate this total "rest" across sectors and add it to the CRS sectoral aid data to make up for the incomplete nature of the latter.

In order to achieve that, we need data from one more table. DAC table 5 comprises aggregate aid and its sectoral allocation, but only by donor; no disaggregation by recipient is possible. So, from this table, I obtain complete data on aggregate aid and its sectoral allocation for each donor and each year. I sum the above CRS data over all recipients, to get it in exactly the same donor-year format. For each donor-year, this leaves me with complete data on aggregate aid and its sectoral allocation (from DAC5), and incomplete data on aggregate and sectoral disbursements (from CRS), as well as a total rest (as calculated above). Consequently, for each donor-year I can calculate the amount of missing data for each sector. So, in addition to the total rest, we now have one rest variable for each sector. As a result, for each donor-year and for each sector I can find the share of the sectoral rest in the total rest. These donor- and year-specific allocations of the total rest across sectors is then applied to the total rest in the original data in recipient-donor-year format. That is, I apply the sectoral rest shares of a given donor-year to the total rests of all recipients to which that donor gives aid in that year. This yields sectoral rest variables in a recipient-donor-year format, which can be added to the CRS sectoral aid variables.

By then summing across donors I obtain sectoral aid variables in a recipient-year format. However, because for some donors insufficient information is available to allocate the total rest across sectors, these sectoral variables still underestimate the total amount of aid received by a recipient in a given year. So, lastly,

[11] The OECD distinguishes between ODA and OA (Official Assistance). OA is ODA directed to countries on part II of the DAC list of aid recipients and to multilateral institutions which primarily benefit Part II aid recipients. I refer to both flows with the term ODA.

I scale up the sectoral variables so that their sum matches aggregate aid disbursements received by the recipient. This strategy to construct the aid data and the details of its implementation are described at length in the appendix (section 7 on page 47).

I isolate the following sectors: education (CRS purpose 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 CRS purpose codes, or similarly the sum of all remaining sectors in DAC5). In addition, following the described procedure, data that splits total health and education aid into four prefix codes is constructed: investment projects (IP), sector programmes (SP), technical cooperation (TC), and other (no mark) (ONM). Definitions are reported in the appendix. The prefix codes are useful because, to some extent, they allow to separate on- and off-budget aid flows. While sector programmes consist of on-budget aid, technical assistance is the most likely of the four to be off-budget. This is especially so because technical assistance may be spent in the donor country, and thus never reach the recipient. Investment projects and other (no mark) should be somewhere in between. So, in a variant of equation (4.1), sectoral aid is disaggregated into its four prefix codes.

Control variables are the age dependency ratio (dependents to working-age population), GDP per capita (thousands of constant 2000 international dollars), urbanisation (urban population, % of total), trade (share of GDP), and the Present Value (PV) of debt (share of GDP). The first four variables come from World Bank (2006c), the debt variable is constructed from a World Bank database (Dikhanov, 2004) updated through to 2004.[12]

Aid and debt relief are potentially correlated, and the latter may also affect social spending. The aid variable that is supposed to capture this is action relating to debt, which consists of debt forgiveness grants, other action on debt (such as service payments to third parties, debt conversions, and debt buybacks), and debt rescheduling (see OECD, 2000b). A drawback is action relating to debt measures the face value of total debt forgiven in a certain year, rather than its present value. Because the average concessionality of debt varies strongly across countries this may be misleading (Depetris Chauvin and Kraay, 2005). Similarly, for 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 present value of current

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

and future reductions in debt service of a rescheduling in the current year.[13]

Hence, instead of including actions relating to debt as a regressor, I omit it and control for the PV of debt, which should pick up the effect of debt relief on social spending. In addition, this measure captures the potentially important effect of the level of government debt on social spending, which a measure of debt relief cannot do.[14] Moreover, action relating to debt would not take into account reductions in debt and debt service owed to creditors that are not included in the DAC, whereas the PV of debt is broader in coverage as it is based on a debtor-based database, namely the World Bank's Debtor Reporting System.

Age dependency is expected to increase social spending, whereas the sign for GDP per capita and trade is ambiguous. High social spending and more specifically the taxes needed to finance this spending may reduce competitiveness in global markets (e.g. by driving up the cost of labour), so more open societies may be more reluctant to spend money on education and health. Globalisation may also erode capital tax bases or make countries more reluctant to tax imports and exports, leading to a race to the bottom which makes it difficult to finance the welfare state. This is the discipline or efficiency effect. However, openness to trade may also increase demand for social spending, which is the so-called compensation effect. Moreover, in many developing countries trade provides an easy tax handle from which spending can be financed. As a result, the effect of trade is ambiguous (Dreher, Sturm, and Ursprung, 2006; Kaufman and Segura-Ubiergo, 2001). So is the effect of urbanisation. On the one hand, it is easier to get public services to people in a more urbanised society. On the other hand, provision costs may increase if a large number of people are concentrated in small areas and – relevant mainly for health spending – the risk of contagion may be higher in cities (see Gerdtham and Jönsson, 2000, for the latter point).

4.2.1 Related studies that use sectoral aid data

A few recent studies have exploited the sectoral aid disaggregation available in the CRS database to analyse a range of issues. However, these studies have typically not taken into account the incomplete nature of the data (Clemens, Radelet, and Bhavnani, 2004; Dreher, Nunnenkamp, and Thiele, 2006; Thiele, Nunnenkamp,

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

[14] That is also the main reason why I prefer to use the PV of debt over the PV of debt relief variable constructed by Depetris Chauvin and Kraay (2005), which was kindly shared by Nicolas Depetris Chauvin.

and Dreher, 2006; Wolf, 2006). Especially in earlier years this is bound to yield misleading conclusions, as CRS coverage becomes very low. In addition, sectoral disbursements before 1990 are often estimated by applying the sectoral allocation found in the commitments data to aggregate disbursements (Clemens, Radelet, and Bhavnani, 2004; Pettersson, 2006). To the extent that sectoral disbursements are determined by past commitments and the allocation of commitments is unstable over time, this again is likely to yield erroneous conclusions. I circumvent this problem by focusing on the post-1990 period, for which disbursements are available.

The studies that do take into account the incomplete nature of the CRS data do so by scaling up CRS aid in a rather simplistic way, applying the sectoral allocation of the available disbursements or commitments data to the complete DAC disbursements or commitments (Michaelowa and Weber, 2006; Pettersson, 2006). This assumes the sectoral aid allocation in the available, incomplete CRS data is undistorted, i.e. it is close to the true sectoral allocation if all donors reported complete data to CRS. This assumption is problematic because donors vary in the sectoral allocation of their aid, and this allocation is likely to be correlated with the share of total aid a donor reports to CRS. In other words, donors that do not report to CRS or only report a small part of their aid, might have a different sectoral allocation than donors that report (almost) all their aid to CRS. Similarly, which sectors a given donor reports might not be random, so the sectoral allocation of the incomplete aid reported by a donor might not accurately reflect the sectoral allocation of that donor's total aid. For these two reasons, scaling up the sectoral aid allocation of a recipient in a given year is likely to introduce measurement error, especially if the reported sectoral CRS data is small with respect to total (complete) aid received.

In contrast to this, the strategy pursued here to construct the sectoral aid variables tries to take into account that donors that report only part, or even none, of their aid to CRS, behave in a different way than donors that report a larger part of their aid, and that the sectoral allocation of a donor's unreported aid might be different from that of the reported aid. For each donor-year, I calculate how aid disbursements "missing" from the CRS database are allocated across sectors. This is done in such a way as to assure the allocation of total aid for each donor mirrors the sectoral allocation in DAC5, which contains complete data. Once a total rest for each donor-year is calculated, and once this rest has been allocated across sectors, the main assumption is that this allocation applies equally to each recipient that receives aid from that donor in that given year. Obviously, for a given donor in a given year it is possible that the sectoral allocation differs across recipients, but no information is available to take this account. Hence, our method is not without flaws, but to the extent that the sectoral allocation of aid is mainly donor-specific, and this allocation does not differ too much across recipients for a given donor, the measurement error we introduce should be limited.

This assumption is most likely to be violated when the rest to be allocated for a donor is small. To see this, suppose a donor has a total rest of 100000 US$ which goes to three recipients (R_1, R_2 and R_3). Suppose, using the methodology described above, we find a quarter of this rest is for health, a quarter for education, and half for general budget assistance. Because the rest is small, it is likely to reflect just three different aid grants (or loans), each of which is given to a different recipient, rather than a larger number of loans and grants that are spread across the three recipients. A possible example might be one grant of 25000 US$ of education aid is given to R_1, the same amount of health aid is given to R_2 under a second grant and 50000 US$ of general aid is given to R_3 under a third grant. However, because I have no further information about how to allocate the donor's rest across the three recipients, I assume each of the recipients receives a third of the education, health and general aid. Hence, for the recipient that receives the education grant (R_1), education aid is underestimated, whereas it is overestimated for the other two recipients. A similar kind of measurement error is introduced for the other two sectors. Consequently, the calculated sectoral aid variables end up being less lumpy and more continuous that the true sectoral aid variables: all three recipients are allocated some education aid, even though in reality only one recipient actually receives education aid. When the total rest is larger it is more likely to consist of a greater number of grants and loans and, similarly, the total rest each recipient receives is more likely to be made up of multiple grants and loans. In such a case, assuming the donor-specific allocation of the rest applies in the same way to all recipients that receive aid from that donor is more likely to be a valid approximation. This also implies measurement error should be smaller for larger sectoral aid values. This is not just because the method to allocate the rest is more accurate when the total rest is large, but also because larger aid values typically comprise some sectoral aid that is recorded in CRS.

Lastly, I scale up the sectoral aid variables for each recipient to make sure the sum of sectoral aid variables matches an aggregate measure of aid received by the recipient in a given year, in a similar way as has been done in previous studies. However, because the data I scale up is based on more extensive information, its sectoral allocation is more likely to reflect the true sectoral allocation of total (complete) aid (see appendix). As before, this is the best that can be done with the available information. Not scaling up the data runs the risk of exaggerating the positive effect of sectoral aid on sectoral spending, if the true relationship between the variables is positive. In general, the most likely outcome of the data construction method is that the sectoral aid variables are somewhat noisy, making it more difficult to find a positive relationship between sectoral aid and spending, if such a relationship exists. Other than that, there is little reason to expect the introduced measurement error to be non-random with respect to sectoral public spending, and to bias the relationship either way.

5 Results

This section discusses preliminary results obtained from estimating versions of (4.1). When linking sectoral aid to sectoral spending other aid flows are controlled for, so as to avoid omitted variable bias. In the equation of public education spending other sector aid includes health aid (I call this variable "other sector aid education"), and similarly, in the equation for public health spending education aid is part of other sector aid (so this variable is labelled "other sector aid health"). Moreover, in both equations commodity aid/general programme assistance and support to NGOs are controlled for. Support to NGOs does not involve the recipient government, but may still elicit a reaction in the recipient's fiscal policy, while part of general aid may be used to finance education and health spending.

The remaining two aid variables (action relating to debt and donor administrative costs) are not included. Though donor administrative costs are allocated by recipient they only affect the donor, and have no influence on recipient behaviour. Actions relating to debt could be used as a measure of debt relief granted, which may affect social spending and could at the same time be correlated with education and health aid. However, as discussed above, we prefer using a broader measure to capture this, namely the PV of debt (as a share of GDP).

Table 2 shows two-way fixed effects results for the full sample for both sectors. Education aid has a very weak and insignificant negative effect, and indeed all aid variables are insignificant. Results for the health sector are somewhat different, as health aid has a positive effect on public health spending. However, the size of the effect is fairly small. Only about a quarter of health aid finances public health

Table 2: Total education and health aid

	Dependent variable	
	Public education spending	Public health spending
Education aid	-.028	
	(.048)	
Health aid		.256***
		(.084)
General aid	.0002	-.012
	(.023)	(.013)
Support to NGOs	-.249	-.129
	(.176)	(.098)
Other sector education	-.011	
	(.011)	
Other sector health		-.013**
		(.006)
Age dependency ratio	-.987	.851
	(1.106)	(.615)
GDP per capita	.343***	.103**
	(.078)	(.046)
Urbanisation	.093***	.075***
	(.022)	(.012)
Trade	-.017***	-.009***
	(.002)	(.001)
PVdebt	-.002**	.00003
	(.001)	(.0006)
Constant	1.248	-1.602**
	(1.362)	(.766)
N	1054	1055
Countries	105	104
R^2 (within)	.165	.127
F-test	8.354***	6.159***
F-test FE	32.179***	34.008***
Hausman	.000***	.000***

Note: Annual data, 1990-2003. All regressions include time dummies, coefficients not reported. The 1990 dummy is dropped. F-test and F-test FE give the F-statistic for the joint significance of the model and the fixed effects, respectively. Hausman reports the p-value of a Hausman test that compares the fixed effects model with a random effects model. Standard errors in brackets. $*$, $**$, and $***$ denote significance at 10, 5 and 1%, respectively.

expenditure (if we assume all health aid is on-budget). Urbanisation and a country's per capita income have a positive effect on both public education and health spending. Conversely, more openness to trade reduces government spending in both sectors. While higher debt decreases education spending, there is no effect on health spending. The results regarding the control variables broadly hold throughout this paper, so they will not be discussed any further in what follows. In both models the joint significance of the variables in the model is not rejected. Similarly, the country-specific fixed effects are jointly significant at a 1% significance level. The Hausman test comfortably rejects the random effects model in favour of the fixed effects model.

So far results are broadly in line with those of past fungibility studies. Education aid has no discernible effect on education spending, while the effect of health aid on public health spending is small. However, as will become clear shortly, the latter result is actually quite a strong one, given that part of health aid flows are off-budget. Despite this, I find an increase in health spending roughly around 0.25% of GDP for every increase in health aid with 1% of GDP.

Table 3 explores this issue in more detail, by disaggregating education and health aid into four prefix codes: investment projects (IP), sector programmes (SP), technical cooperation (TC) and other (no mark) (ONM). Recall, a priori technical cooperation is expected to be the most off-budget aid flow of the four. Conversely, all sector programme aid should be on-budget. Moreover, a priori one would expect TC to be the least fungible of all four prefix codes. Very few developing countries spend domestic resources on the type of goods and services financed by technical assistance. Hence, when donors step in to finance such expenditure, this aid is likely to be additional: recipients cannot cut back spending from own resources on goods and services that are similar in nature as those financed by technical assistance because they simply do not undertake such spending.[15] Conversely, since SP concerns aid given in cash or in kind, it is much more likely to be fungible.

In the education sector column of table 3 education SP is the only prefix variable that is positive and significant, and its coefficient is close to and not significantly different from 1. In contrast, the effect of education TC is zero, and the coefficients for education IP and ONM are also insignificant. Results for health aid, disaggregated into prefix codes, are similar. Health SP has a positive effect,

[15]Gramlich (1977) makes the general argument that if government spending is not homogeneous it may becomes physically impossible to displace expenditure from own sources as a response to grants, because such expenditure may not be there in the first place. As such, heterogeneity in government spending would tend to limit fungibility.

Table 3: Education and health aid, disaggregated into prefix codes

	Dependent variable	
	Public education spending	Public health spending
Education aid IP	.096	
	(.171)	
Education aid SP	.986**	
	(.445)	
Education aid TC	-.057	
	(.059)	
Education aid ONM	.032	
	(.118)	
Health aid IP		.252*
		(.142)
Health aid SP		.841***
		(.234)
Health aid TC		.047
		(.141)
Health aid ONM		.310**
		(.140)
General aid	-.0008	-.011
	(.023)	(.013)
Support to NGOs	-.248	-.117
	(.176)	(.098)
Other sector education	-.011	
	(.011)	
Other sector health		-.015**
		(.006)
Age dependency ratio	-1.453	.836
	(1.123)	(.623)
GDP per capita	.360***	.110**
	(.079)	(.046)
Urbanisation	.097***	.075***
	(.022)	(.012)
Trade	-.016***	-.009***
	(.002)	(.001)
PVdebt	-.003**	.00003
	(.001)	(.0006)
Constant	1.372	-1.610**
	(1.365)	(.772)
N	1053	1055
Countries	105	104
R^2 (within)	.17	.137
F-test	7.58***	5.875***
F-test FE	31.998***	33.986***
Hausman	.000***	.000***

Note: Annual data, 1990-2003. All regressions include time dummies, coefficients not reported. The 1990 dummy is dropped. F-test and F-test FE give the F-statistic for the joint significance of the model and the fixed effects, respectively. Hausman reports the p-value of a Hausman test that compares the fixed effects model with a random effects model. Standard errors in brackets. $*$, $**$, and $***$ denote significance at 10, 5 and 1%, respectively.

which is insignificantly different from 1, though the coefficient is a bit smaller than for the education sector. The coefficient for health TC is insignificant. In addition, health IP and ONM increase health spending, though to a lesser extent than health SP.

As an initial check of the robustness of results, table 4 adds additional control variables which have been included in some past studies as potential determinants of public education and health spending. The source for all variables is again World Bank (2006c). Note that, when the birth rate (crude, per 1000 people) is included in columns (1) and (3), there is a massive drop in the number of observations. Hence, I also show results excluding the birth rate, which only increases public health spending but has no significant effect on public education spending. In general, in the reduced sample the effects of most prefix codes become stronger, but they are also estimated less precisely. In the more complete sample, results are very similar to those found before. For both sectors SP aid has a strong positive effect, which is insignificantly different from one. Of the additional control variables only the growth of real per capita GDP (constant Local Currency Units) has a consistent – negative – effect on spending in both sectors, so I include this variable in the model for the remainder of the paper. Female labour force participation reduces public health spending, which is unexpected, and has no effect on public education spending, while population (in millions) affects neither.

The PV of debt may not adequately capture immediate financing constraints in the current year imposed by large amounts of debt service. To check whether this affects results, debt service (as a share of GDP) is added to the model in table 5. Indeed, debt service is found to have an additional negative effect on public education and health spending, over and above the negative effect of the PV of debt. Columns (2) and (4) show similar conclusions are reached when the PV of debt is replaced by public and publicly guaranteed debt (as a share of GDP): higher debt again reduces public education but not public health spending while debt service reduces public spending in both sectors.[16] As far as the aid prefix codes are concerned, results do not change.

Results hold when low values for "scaling" are excluded. Recall, in the final step of the data construction I scale the sectoral aid variables so that their sum equals an aggregate measures of aid disbursements received. Hence, the scaling variable is calculated as the ratio of the sum of the sectoral aid variables and aggregate disbursements, as described in more detail in the appendix (table 8 in the

[16]Both debt variables are taken from World Bank (2006b)

Table 4: Extra control variables

	Dependent variable			
	Public education spending		Public health spending	
	(1)	(2)	(3)	(4)
Education aid IP	.534**	.077		
	(.249)	(.168)		
Education aid SP	1.419**	1.210***		
	(.590)	(.440)		
Education aid TC	-.046	-.063		
	(.091)	(.058)		
Education aid ONM	.259	.044		
	(.198)	(.116)		
Health aid IP			.055	.215
			(.202)	(.138)
Health aid SP			1.352***	.937***
			(.445)	(.228)
Health aid TC			.449**	.039
			(.188)	(.137)
Health aid ONM			.907***	.384***
			(.200)	(.136)
General aid	.012	-.007	.008	-.013
	(.045)	(.022)	(.023)	(.012)
Support to NGOs	-.536*	-.277	-.295*	-.127
	(.288)	(.173)	(.150)	(.095)
Other sector education	-.004	-.008		
	(.014)	(.011)		
Other sector health			-.027***	-.014**
			(.008)	(.006)
Age dependency ratio	-2.306	-1.860*	.189	.632
	(1.659)	(1.107)	(.840)	(.606)
GDP per capita	.279***	.354***	.142***	.140***
	(.097)	(.079)	(.053)	(.045)
Urbanisation	.122***	.091***	.086***	.069***
	(.029)	(.022)	(.015)	(.012)
Trade	-.013***	-.014***	-.006***	-.008***
	(.003)	(.002)	(.002)	(.001)
PVdebt	-.002	-.003***	-.00009	-.0002
	(.001)	(.001)	(.0007)	(.0006)
Birth rate	.032		.063***	
	(.042)		(.022)	
Population	-.006	-.004	-.003	-.002
	(.006)	(.005)	(.003)	(.003)
Female labour force part	-.027	-.020	-.043**	-.047***
	(.041)	(.031)	(.022)	(.017)
Real GDP per capita growth	-.037***	-.036***	-.023***	-.025***
	(.009)	(.006)	(.005)	(.003)
Constant	.925	2.780	-2.212	.522
	(2.584)	(1.781)	(1.387)	(1.003)
N	588	1052	589	1055
Countries	105	105	103	104
R^2 (within)	.216	.205	.26	.192
F-test	4.303***	8.484***	5.543***	7.843***
F-test FE	14.861***	32.582***	21.208***	34.731***
Hausman	.000***	.000***	.000***	.000***

Note: Annual data, 1990-2003. All regressions include time dummies, coefficients not reported. The 1990 dummy is dropped. F-test and F-test FE give the F-statistic for the joint significance of the model and the fixed effects, respectively. Hausman reports the p-value of a Hausman test that compares the fixed effects model with a random effects model. Standard errors in brackets. $*$, $**$, and $***$ denote significance at 10, 5 and 1%, respectively.

Table 5: Alternative debt variables

	Dependent variable			
	Public education spending		Public health spending	
	(1)	(2)	(3)	(4)
Education aid IP	.124 (.164)	.131 (.164)		
Education aid SP	1.299*** (.430)	1.290*** (.429)		
Education aid TC	-.041 (.059)	-.052 (.059)		
Education aid ONM	.022 (.113)	.024 (.113)		
Health aid IP			.241* (.137)	.225 (.137)
Health aid SP			.894*** (.226)	.895*** (.226)
Health aid TC			.032 (.135)	.046 (.136)
Health aid ONM			.390*** (.135)	.412*** (.135)
General aid	.018 (.022)	.023 (.023)	-.004 (.013)	-.002 (.013)
Support to NGOs	-.345** (.170)	-.311* (.170)	-.177* (.094)	-.165* (.094)
Other sector education	-.005 (.011)	-.001 (.011)		
Other sector health			-.015** (.006)	-.014** (.006)
Age dependency ratio	-1.137 (1.080)	-1.021 (1.078)	1.097* (.599)	1.160* (.600)
GDP per capita	.305*** (.079)	.303*** (.078)	.122*** (.044)	.124*** (.044)
Urbanisation	.093*** (.021)	.091*** (.021)	.070*** (.012)	.069*** (.012)
Trade	-.012*** (.002)	-.012*** (.002)	-.006*** (.001)	-.006*** (.001)
PVdebt	-.003*** (.001)		-.0002 (.0006)	
Real GDP per capita growth	-.041*** (.006)	-.041*** (.006)	-.026*** (.003)	-.026*** (.003)
Debt service	-.061*** (.014)	-.056*** (.014)	-.023*** (.008)	-.021** (.008)
Debt GDF		-.004*** (.001)		-.0008 (.0006)
Constant	1.383 (1.311)	1.420 (1.308)	-1.729** (.742)	-1.724** (.742)
N	1038	1038	1044	1044
Countries	103	103	103	103
R^2 (within)	.227	.23	.199	.201
F-test	9.848***	10.066***	8.435***	8.512***
F-test FE	34.405***	34.318***	36.488***	36.533***
Hausman	.000***	.000***	.000***	.000***

Note: Annual data, 1990-2003. All regressions include time dummies, coefficients not reported. The 1990 dummy is dropped. F-test and F-test FE give the F-statistic for the joint significance of the model and the fixed effects, respectively. Hausman reports the p-value of a Hausman test that compares the fixed effects model with a random effects model. Standard errors in brackets. $*$, $**$, and $***$ denote significance at 10, 5 and 1%, respectively.

appendix also contains some key statistics and percentiles for the scaling variable). So, low values of scaling reflect a lack of information: when scaling is low the available sectoral aid measures only form a small fraction of aggregate disbursements received, so the scaling in the last step of the data construction in such cases is liable to induce a greater amount of measurement error. However, dropping observations with scaling values up until 0.5, which excludes a bit more than 10% of the sample, does not change the results markedly. Results are also robust to the exclusion of outliers in the aid prefix and public spending variables based on graphical eyeballing (results not reported). The downside of a graphical way to identify outliers is that is it necessarily ad hoc and somewhat arbitrary. However, because most observations for the prefix codes are bunched tightly around small values, it is difficult to use a mechanical rule to identify outliers. Moreover, given the way the data has been constructed larger values for the aid prefix codes are less likely to be mismeasured than smaller values, as discussed in section 4.2.1. As an alternative way to check the influence of outliers I estimate a double-log version of equation (4.1):

$$\begin{aligned} LN(1+SSP_{i,t}) = \alpha + \beta LN(1+SAID_{i,t}) + \sum_{k=1}^{K} \gamma_k LN(1+A_{i,t}^{k}) + \\ \sum_{l=1}^{L-1} \delta_l LN(1+X_{i,t}^{l}) + X_{i,t}^{L} + \eta_i + \epsilon_{i,t} \end{aligned} \tag{5.1}$$

where all variables are defined as above, and LN refers to the natural logarithm. $X_{i,t}^{L}$ is the growth of real GDP per capita, of which I do not take the natural logarithm. Note the natural logarithm of one plus each variable is taken for the other variables.[17] As already mentioned, for many variables, and especially for the aid prefix codes, a lot of observations are between zero and one. Indeed, the majority of observations for the aid prefix variables are close to zero. If we would simply take the log these values would turn negative and be more dispersed than before taking logs, whereas higher aid values are pressed closer together after taking the log. In the regressions this would give undue weight to the observations that barely exceed zero, which are more likely to be mismeasured than the higher values. Hence, we would needlessly be emphasising the most uninformative part

[17] Also, before taking the logarithm, real GDP per capita is expressed in constant 2000 international dollars instead of *thousands* of constant 2000 international dollars.

of the data. Consequently, I add one before taking natural logarithms. This has the advantage that observations with a zero value remain zero after the transformation, whereas strictly positive values remain strictly positive.

β now estimates an elasticity, and the marginal effect depends on the level of sectoral spending and aid:

$$ME_\beta = \frac{\partial SSP}{\partial SAID} = \beta(\frac{1+SSP}{1+SAID}) \tag{5.2}$$

This shows a downside of this specification, namely that ME_β is restricted to fall with $\frac{1+SAID}{1+SSP}$. The discussion is section 2 (see figure 2), however, suggests that the effect of aid on spending should be stronger (and fungibility smaller) when earmarked sectoral aid flows are large in comparison to sectoral spending. I calculate marginal effects at the means of the variables, taken over all observations that are included in the regression.

Results are reported in table 6, while table 7 shows the calculated marginal effects. As before, the marginal effect of SP aid is close to 1 in both sectors, whereas the other prefix codes, except ONM in the health sector, are insignificant. Results for the control variables are also broadly unchanged.

5.1 Endogeneity

So far our findings suggest the effect of earmarked sectoral aid on public sectoral spending is pulled down by technical assistance, which is likely to be mostly off-budget and is a large component of sectoral aid. Hence, a failure to distinguish between on- and off-budget would lead to the erroneous conclusion that most aid is fungible. In contrast, the effects of education and health sector programmes, which are almost certainly on-budget, are close to one. Hence, results in this paper suggest fungibility is much less prevalent than past studies have shown it to be. While the effect is typically stronger for education than for health, the pattern of coefficients is remarkably similar for both sectors.

To what extent do these results reflect the causal impact of sectoral aid on sectoral spending? For endogeneity to drive our main result, it should lead to a positive bias in the coefficient of SP. This is not entirely unlikely. Developing countries with high public education and health spending have, in all likelihood, developed full-blown sectoral spending programs for education and health. Such recipients have most likely already identified priority areas within each sector, and reduced the wastefulness of spending. It is not unlikely donors choose to support such

Table 6: Double-log model, aid disaggregated into prefix codes

	Dependent variable	
	Public education expenditure	Public health expenditure
Education aid IP	.016	
	(.055)	
Education aid SP	.249**	
	(.113)	
Education aid TC	-.019	
	(.037)	
Education aid ONM	.0006	
	(.047)	
Health aid IP		.088
		(.060)
Health aid SP		.438***
		(.097)
Health aid TC		.030
		(.066)
Health aid ONM		.152**
		(.062)
General aid	-.012	-.007
	(.019)	(.017)
Support to NGOs	-.129**	-.120**
	(.061)	(.054)
Other sector education	-.008	
	(.026)	
Other sector health		-.051**
		(.024)
Age dependency ratio	-.259	.663*
	(.404)	(.354)
GDP per capita	.232***	.186***
	(.060)	(.056)
Urbanisation	.496***	.634***
	(.161)	(.144)
Trade	-.122***	-.079**
	(.039)	(.034)
PVdebt	-.060***	-.031**
	(.018)	(.016)
Real GDP per capita growth	-.006***	-.006***
	(.001)	(.001)
Constant	-1.189	-2.676***
	(.797)	(.726)
N	1052	1055
Countries	105	104
R^2 (within)	.174	.178
F-test	7.451***	7.725***
F-test FE	30.275***	35.799***
Hausman	.002***	.000***

Note: Annual data, 1990-2003. The natural logarithm of 1 plus the variable is taken for all variables except real GDP per capita growth. All regressions include time dummies, coefficients not reported. The 1990 dummy is dropped. F-test and F-test FE give the F-statistic for the joint significance of the model and the fixed effects, respectively. Hausman reports the p-value of a Hausman test that compares the fixed effects model with a random effects model. Standard errors in brackets. $*$, $**$, and $***$ denote significance at 10, 5 and 1%, respectively.

Table 7: Marginal effects derived from table 6

	Dependent variable	
	Public education expenditure	Public health expenditure
Investment projects	.070	.227
Sector programmes	1.191**	1.209***
Technical cooperation	-.053	.073
Other (no mark)	.002	.400**

Note: the columns show the marginal effects corresponding to the estimated elasticities in the same column of table 6. $*$, $**$, and $***$ denote significance of the estimated elasticity at 10, 5 and 1%, respectively.

countries with sector programmes, leaving it up to the recipient to decide how best to spend the funds within the sector. On the other hand, countries with low levels of public education and health spending are likely to have a less developed sectoral spending framework, and the quality of their spending is probably lower. In such cases donors may feel uneasy to grant aid in the form of sector programmes. Rather, donors might seek to help establish a framework for spending, and improve the quality of spending, which can best be achieved through technical cooperation (this would induce a negative bias in the TC coefficients). Hence, our results may suffer from endogeneity, not all of which may be picked up by the country-specific fixed effects: education and health policies might change over time within countries, perhaps because of elections or regime changes, and donors' aid allocation may respond to this. Think, for instance, of a country that announces a policy of universal primary education, to which donors may respond by increasing SP aid to the education sector. Hence, the task ahead is to isolate the exogenous effects of the aid prefix variables, especially SP.

6 Conclusion

This paper presents new evidence to shed light on the thorny issue of foreign aid fungibility. Theory is not conclusive about the expected extent of fungibility. Though full fungibility arises as the natural response of an optimising government to an earmarked aid inflow, donors may be able to curtail the extent of fungibility through a combination of earmarking, choice of modality, monitoring, and effective conditionality. From a review of the literature I conclude fiscal response models, which are the predominant tool with which fungibility has been studied to date, are based on weak theoretical foundations and are flawed in their empirical application. As a result, few conclusions can be drawn based on the results of these studies. Fungibility studies, though not without problems themselves, suggest aid is, to a large extent, fungible.

To reassess this evidence, I construct measures of aid earmarked for education and health, and link these aid measures to public spending in the same sectors. Using information on the sectoral allocation of aid provides a marked improvement on most previous studies, which have no or incomplete information on the purpose for which aid is given. However, this is insufficient to overturn the conclusions reached by fungibility studies: education aid has no effect on education spending, and though the effect of health aid on public health spending is positive it is also limited in size.

This paper further argues a failure to distinguish between on- and off-budget aid exaggerates the extent of fungibility, thus undermining the effect of sectoral aid on sectoral spending. Empirical evidence broadly supports this argument. Sectoral programmes, which should be wholly on-budget, have a strong positive effect (close to one) on sectoral spending. Technical cooperation, the bulk of which is probably off-budget, has no discernible effect on public spending. Our results, though falling short of providing a razor-sharp estimate of the extent of fungibility, thus strongly indicate foreign aid is much less fungible than generally assumed to be the case. These results hold up in a variety of specifications and are remarkably similar for both sectors. However, results may be driven by the endogenous response of donors to changes in a recipient's education and health spending. Hence, I aim to take the empirical analysis further in order to examine whether the results obtained indeed reflect a causal effect.

References

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

BINH, T.-N., AND M. MCGILLIVRAY (1993): "Foreign aid, taxes, and public investment: a comment," *Journal of Development Economics*, 41(1), 173–176.

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).

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 (1999): "What does aid to Africa finance," World Bank, Policy Research Working Paper 2092.

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

DREHER, A., P. NUNNENKAMP, AND R. THIELE (2006): "Does aid for education educate children? Evidence from panel data," Kiel Institute for the World Economy, Working Paper No. 1290.

DREHER, A., J.-E. STURM, AND H. W. URSPRUNG (2006): "The impact of globalization on the composition of government expenditures: evidence from panel data," CESifo, Working Paper No. 1755.

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.

FRIEDMAN, M. (1964): "Foreign economic aid: means and objectives," in *The United States and the developing economies: an introduction to a current issue* *of public policy*, ed. by G. Ranis, pp. 24–38. W.W. Norton & Company, New York, United States.

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.

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.

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

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

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.

KHAN, H. A., AND E. HOSHINO (1992): "Impact of foreign aid on the fiscal behavior of LDC governments," *World Development*, 20(10), 1481–1488.

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.

LORA, E., AND M. OLIVERA (2006): “Public debt and social expenditure: friends or foes,” Inter-American Development Bank, Research Department, Working Paper 563.

MAVROTAS, G. (2002): “Foreign aid and fiscal response: does aid disaggregation matter,” *Weltwirtschaftliches Archiv*, 138(3), 534–559.

MAVROTAS, G., AND B. OUATTARA (2006): “Public fiscal behaviour and aid heterogeneity in aid-recipient economies,” *Journal of Developing Areas*, 39(2), 1–15.

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. (1994): “The impact of foreign aid on the fiscal behavior of Asian LDC governments: a comment on Khan and Hoshino (1992),” *World Development*, 22(12), 2015–2017.

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

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. (1978): “A method for estimating the effect of a subsidy on the receiver’s resource constraint: with an application to U.S. local governments 1964-1971,” *Journal of Public Economics*, 10(1), 25–44.

——— (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).

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

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

——— (2002): “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.

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.

PETTERSSON, J. (2006): “Foreign sectoral aid fungibility, growth, and poverty reduction,” Stockholm University, Department of Economics, mimeo (May 2006 version).

STEIN, R. M., M. ISHIMATSU, AND R. J. STOLL (1985): “The fiscal impact of the US military assistance program, 1967-1976,” *Western Political Quarterly*, 38(1), 27–43.

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.

THIELE, R., P. NUNNENKAMP, AND A. DREHER (2006): “Sectoral aid priorities: are donors really doing their best to achieve the Millennium Development Goals,” Kiel Institute for the World Economy, Working Paper No. 1266.

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

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

——— (2006): "The Millennium Development Goals report 2006," .

WHITE, H. (1994): "Foreign aid, taxes, and public investment: a further comment," *Journal of Development Economics*, 45(1), 155–163.

WOLF, S. (2006): "Aid and public expenditure in Africa: does aid increase the efficiency of public service delivery," Paper prepared for WIDER conference on aid: principles, policies and performance, 16-17 June 2006.

WORLD BANK (1998): "Assessing aid: what works, what doesn't and why," .

——— (2006a): "Edstats," .

——— (2006b): "Global Development Finance," .

——— (2006c): "World Development Indicators," .

ZAHARIADIS, N., R. TRAVIS, AND P. F. DIEHL (1990): "Military substitution effects from foreign economic aid: buying guns with foreign butter," *Social Science Quarterly*, 71(4), 774–785.

ZAMPELLI, E. M. (1986): "Resource fungibility, the flypaper effect, and the expenditure impact of grants-in-aid," *Review of Economics and Statistics*, 68(1), 345–362.

7 Appendix: construction of the aid data

As already discussed in the main text, the OECD's Creditor Reporting System (CRS) contains aggregate disbursements, but also disaggregates aid by purpose. Gross disbursements are available for 1990-2004. This data can be obtained in a recipient-donor-year format, i.e. for each year it shows how much aid is transferred from a given donor to a given recipient. Unfortunately, the CRS disbursements reported by some multilateral and bilateral donors are incomplete or even nonexistent for some years, which is why they need to be complemented with additional information. All aid data in the OECD's International Development Statistics online is expressed in millions of US$.

CRS data on the following purposes is downloaded: education (purpose 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 purpose codes).[18] Education and health disbursements are further disaggregated into four prefix codes: investment projects (IP), sector programmes (SP), technical cooperation (TC), and other (no mark) (ONM).

The prefix codes are useful because, to some extent, they allow to separate on- and off-budget aid flows. Full 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 the 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. Sector programme aid comprises contributions to carry out wide-ranging development plans in a defined sector. Assistance is made available in cash or in kind, 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 to finance the supply of goods and services in support of such schemes. This includes investment-related technical cooperation, which is the financing of services (e.g. planners, engineers,

[18]Whereas the OECD makes a distinction between sector allocable and non-sector allocable aid, for reasons of convenience I refer to the underlying series in both categories as sectors (or purposes). 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. Whenever it is necessary to make the distinction between sector allocable and non-sector allocable aid, I will do so explicitly by using these terms. In all other instances, "sectors" can be interpreted more broadly as referring to the series in both categories.

technicians,...) with the primary purpose of contributing to the design and/or implementation of a project or programme aiming to increase the physical capital stock. Other (no mark) is the residual category. Hence, while sector programmes consist of on-budget aid, technical assistance is the most likely of the four to be off-budget, especially since technical assistance may be spent in the donor country. Investment projects and other (no mark) should be somewhere in between.

At this stage it is important to note the database does not record zeros. When no aid is given for a certain sector the observation is simply missing, so in general it is difficult to tell whether an observation is missing because zero aid is given or because existing aid flows are not reported. So, whenever total education or health disbursements are available, which is the case when at least one of the four prefix codes is available, missing values for the other prefix codes are set to zero. Similarly, whenever aggregate disbursements are available, missing observations for the sectoral aid variables, as well as the 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 very small rounding errors.

I also download CRS data on grants and loans separately, 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, apart from some extremely small rounding errors (occurring in the fifth digit after zero). I eliminate these rounding errors by recalculating aggregate CRS disbursements as the sum of CRS grants and loans extended.

The aggregate and sectoral disbursements thus obtained from CRS in a recipient-donor-year format form the backbone of our data construction. From here on I refer to this dataset as CRS-RDY (RDY stands for recipient-donor-year). Because these aid measures are incomplete I attempt to improve on them, employing data from DAC table 2a. This table contains complete data on grants and loans extended, again in a recipient-donor-year format, and is henceforth referred to as DAC2a-RDY. I exclude recipients in this table that are not available in CRS. Conversely, for Serbia CRS data is available but DAC data is not, so Serbia is also dropped. In addition, I only select donors that are also available in DAC table 5, for reasons that will become clear in a bit. Missing values for loans are set to zero when grants are observed, and vice versa. Total disbursements are then calculated as the sum of loans extended and grants.

We 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, we obtain a “rest” for aggregate disbursements, i.e. the amount of aid disbursements from a donor to a recipient in a given year that is not recorded in CRS, or the amount of disbursed aid that is “missing” from the CRS database. Our aim is to allocate this rest across sectors.

There are quite a few values for which this rest is negative (DAC2a disbursements are smaller than CRS disbursements). In the majority of cases the differences are very small, though there are a few observations for which the absolute value of the difference is large. Consequently, I replace DAC grants (loans) by the CRS amount in all cases where CRS grants (loans) exceed DAC grants (loans). I then recalculate DAC2a disbursements as the sum of DAC2a grants and loans, and recalculate the rest variable. However, I make sure the DAC value is not replaced by the CRS value when the DAC value is negative and the CRS value is zero. But when the DAC value is negative and the CRS value is non-zero then the former *is* replaced by the latter.

The rationale for these adjustments is that it is very unlikely that aid is reported even though it never actually took place. It is far more likely actual aid is underreported, i.e. it is more likely DAC figures are missing something when they are exceeded by CRS figures, even though they are supposed to be complete. Obviously, it might also be the case that negative amounts of aid go unreported in CRS, and this is what causes the CRS figure to exceed the DAC figure. However, this is less likely, mainly because it is not clear what such negative amounts of aid mean (even though there are a few examples of negative aid in the dataset) and, as a result, negative amounts of aid are expected to be rare.

Applying this rationale consistently is also what leads me not to replace the DAC value by the CRS value when the former is negative and the latter is zero. A CRS value of zero means no aid is reported to CRS, while the negative value for DAC implies there was some aid, albeit negative. As a result, there are a few observations (136 observations out of a total of 43216) with a negative rest. The situation where DAC aid is negative and CRS aid is non-zero is more tricky. On the one hand, the DAC 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 DAC amount by the non-zero (and always positive) CRS amount. Because there are only a few such cases (9 for grants, 17 for loans, out of a total of 43216 observations), this choice is unlikely to have a large impact on the data.

There are some recipient-donor-year observations for which CRS data is available but DAC data is unavailable (1230 out of a total of 43216), which arises because some donors that are included in CRS are not available in DAC. These donors are the Global Fund to Fight Aids, Tuberculosis and Malaria (GFATM), United Nations Population Fund (UNFPA), and the Joint United Nations Programme on HIV and AIDS (UNAIDS). For these observations no rest can be calculated. However, I do not delete these observations from the CRS database, they are simply treated as having zero rest. Conversely, when observations are available from DAC but missing from CRS, all CRS variables are set to zero so that the complete DAC value is recorded as rest.

Having calculated a total rest for each recipient-donor-year observation, I collapse this dataset by summing across recipients, yielding a rest variable for aggregate disbursements in a donor-year format. In this collapsed dataset (which I refer to by using abbreviation DY, for donor-year) the rest variable is never negative. The reason for collapsing the dataset is that now, with data from one more table, it becomes possible to allocate the total rest for each donor-year across sectors. While some of the rest values are negative in the RDY dataset (113 out of a total of 43216), rest-DY is always positive.

In order to allocate the total rest across sectors one more piece of information, which comes from DAC table 5, is needed. DAC5 contains a sectoral disaggregation of total Official Development Assistance (ODA), but in a donor-year format, not by recipient (call this DAC5-DY). Missing observations for the sectoral aid variables are turned to zero whenever total ODA is available. As before, other sector aid is calculated as the sum of all sectors that are not isolated individually. A problem here is that total ODA is not always equal to the sum of the underlying sectoral aid variables. Four observations show up with large discrepancies: France, 1997; AsDF (Asian Development Fund), 2002; IDB (Inter-American Development Bank) Special Fund, 1996; and AsDF, 1996 (this is also the minimum observation for total ODA).

For France, 1997, and AsDF, 2002, total ODA exceeds the sum of the sectoral aid variables. In both cases this is because total sector allocable aid exceeds the sum of its underlying series.[19] Hence, for both observations I scale up all sector

[19]Recall sector allocable aid is made up of aid for social infrastructure and services, economic

allocable series so that their sum matches total sector allocable aid. This means education and health aid are scaled up, but also other sector allocable series which are part of other sector aid. Hence, 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 negligible discrepancies total ODA is recalculated as the sum of the sectoral aid variables. For AsDF, 1996, and IDB Special Fund, 1996, the sectoral sum exceeds total ODA, so these observations are also taken care of in this way.

Lastly, from DAC5 I also download data that splits total health and education aid into the four prefix codes, again in a donor-year format. I scale up the education and health prefix codes for France, 1997, and AsDF, 2002, 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. For education there is one observation (Netherlands, 2003) for which the education total exceeds the sum of the prefix codes; for health there are three such observations (Luxembourg, 1993; Netherlands, 2003; Norway, 1992). So, for these observations I scale up the prefix codes so that their sum matches the sector total. For all observations I then recalculate education and health totals as the sum of their prefix codes. This takes care of the one observation for which the sum of the prefix codes exceeds the sectoral total (Netherlands, 1992, for both education and health). It also sorts out the many observations for which there are extremely small discrepancies, again probably due to rounding errors. Because this leads to changes in the values of education and health aid, I recalculate total ODA in DAC5-DY as the sum of the underlying sectors to make sure both are consistent again.

I now have, in donor-year format, (supposedly) complete aid data disaggregated by sector from DAC5 (DAC5-DY), and incomplete aid data disaggregated by sector from the collapsed CRS dataset (CRS-DY). The plan is to calculate sectoral rest variables for each donor-year using this data, and use these to allocate the total rest across sectors. Going back to the data in recipient-donor-year format (rest-RDY) this donor- and year-specific sectoral allocation of the total rest is then applied to all recipients that receive aid from the relevant donor in a given year.

There is, however, one problem that needs to be solved before proceeding. The

infrastructure and services, production sectors, and multisector/crosscutting aid.

sectoral rest variables must be calculated from DAC5 data, whereas the total rest is calculated based on DAC2a data. There are two possible differences in the donor-year data derived form these tables. Firstly, the set of recipients may differ. In the DAC2a dataset I have excluded several recipients because they are not available in CRS, whereas aid given by each donor in DAC5 includes aid given to *all* part I recipients but excludes part II recipients (as OA is omitted in DAC5). Secondly, DAC5 contains a mix of commitments and disbursements. Donors choose to report commitments or disbursements, and no information is available to trace who reports what.

As a consequence, before using the sectoral aid data in DAC5 all aid variables from this table (including the education and health prefix codes) are scaled by the ratio of aggregate DAC2a-DY disbursements to total DAC5-DY ODA, so that the sectoral aid variables from DAC5-DY sum to DAC2a-DY aggregate disbursements. 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 given to a somewhat different set of recipients. The correlation between DAC5-DY ODA and DAC2a-DY disbursements is very high (0.9036) and typically DAC5-DY ODA exceeds DAC2a-DY disbursements. Most likely, the assumption made here is not inaccurate to the extent that it would greatly influence results. The main effect is probably an increase in random measurement error. A few observations of positive DAC5-DY ODA are scaled to zero because DAC2a-DY disbursements are zero (these observations have no aggregate disbursements rest that needs to be allocated anyway).

Note the scaling itself does not change anything about the sectoral allocation of total ODA in DAC5, this information is completely reserved. It does, however, bring the sectoral aid data from DAC5 on a comparable scale to the DAC2a aggregate disbursements. This allows for a calculation of sectoral rests that is more consistent with that of the calculation of the total rest, which is based on DAC2a aggregate disbursements.

If, after the scaling, sectoral values for CRS-DY exceed those for DAC5-DY, the DAC value is replaced by the CRS one. I first do this for the prefix codes, and recalculate total education and health aid as the sum of the prefix codes. I then repeat this strategy for the sectoral aid variables, including education and health aid. At this stage the only changes for education and health totals occur for observations for which there is no further 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 DAC value is not replaced by the CRS value when the DAC value is negative and the CRS value is zero. However, when the DAC 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 (0.9943) of the sum of the sectoral DAC variables (after scaling and replacement) with aggregate DAC2a-DY disbursements.

The total rest in donor-year format (rest-DY) is calculated as the sum of these sectoral aid variables minus aggregate CRS disbursements. The correlation with the collapsed rest variable that was calculated earlier in the recipient-donor-year dataset is 0.9687. The sectoral rest variables, calculated by subtracting CRS from DAC sectoral aid variables, sum to this total rest, and rest variables for the prefix codes sum to the total rest for education and health, apart from extremely small rounding errors (typically in the fifth digit after zero). Whenever the CRS value is not available, the full DAC value is recorded as rest, as before.

Two sectoral rest variables are negative. Finland, 1991, has a negative rest for health IP. For this observation I set the health prefix code disaggregation to missing. This means that the health rest can no longer be allocated across prefix codes, but the education rest can still be allocated across prefix codes and the total rest can still be allocated across sectors. UK, 1996, has a negative rest for action relating to debt. Because this observation has a large total rest it would be a shame to lose it. Moreover, the absolute value of the negative action relating to debt rest is less than 0.1% of the total rest. Hence, measurement error is probably best minimised by setting the action relating to debt rest to zero for this observation and recalculating total disbursements (rather than setting this observation to missing). I then recalculate the total rest for this observation.

Now we are ready to calculate the share of the sectoral rest variables in the total rest, and also the share of the prefix rest variables in the total rest for education and health.[20] Recall, all this is still being done in a donor-year format. These donor- and year-specific allocations of the total rest variable (rest-DY) are then applied to the total rest calculated in a recipient-donor-year format (rest-RDY). That is, we assume the allocation of a donor's rest variable is the same across all recipients for which this donor has a rest. Education and health rests are allocated in a similar way by further multiplying these rest variables with the shares of the prefix rest

[20] Again, there are tiny rounding errors, so the actual denominators used are the sum of all sectoral rests (instead of the total rest), and the sum of all prefix code rests for education and health (instead of total education and health rests), respectively.

variables in the total education and health rests.

I then sum the sectoral CRS data (CRS-RDY) with the calculated sectoral rests in the recipient-donor-year database, and likewise for the education and health prefix codes. Obviously, for observations that have no allocation of the rest, this sum is equal to the CRS sectoral amounts. For some observations insufficient information is available from DAC5 to allow us to allocate the rest. As a result, the sum of the newly calculated sectoral variables does not necessarily equal DAC2a-RDY total disbursements. Similarly, education and health prefix codes do not always sum to the sectoral total, because for some donors insufficient information is available to allocate the education and health rest across prefix codes.[21]

I collapse this recipient-donor-year dataset again, this time by summing across donors, to get the data in a recipient-year format. In this final recipient-year dataset there are observations for which both DAC2a-RY and CRS-RY disbursements are zero. The reason why these observations are zero rather than missing (as we would expect) is that Stata turns missing values into zero when collapsing data (the described data construction is carried out in Stata version 9.2). All aid variables, including prefix codes, are set to missing for these observations. Moreover, there are seven observations (out of a total of 2416 observations) with non-zero DAC2a-RY aggregate disbursements but zeros for all sectoral variables. Since for these observations there is no information whatsoever on how disbursements are allocated, all variables are set to missing. Similarly, there is one observation with zeros for all health prefix codes, but a non-zero health total. For this observation health prefix codes are turned to missing, given that there is no information on the allocation of the health total across prefix codes.

As before the collapse, when I sum the sectoral data I do not always get a number that equals the DAC2a disbursements (DAC2a-RY), and, similarly the sum of the prefix codes does not always equal total education and health aid. So, the last thing I do is scale up the prefix codes in this recipient-year format so that their sum equals total education and health aid, and scale up the sectoral variables so that their sum matches a plausible aggregate measure of aid disbursements the recipient has received in each given year.

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

[21] Conversely, there are also observations for which CRS total disbursements are zero, but a DAC total is available which has been allocated across the different sectors. Hence, for such recipients with no sectoral CRS data, the sectoral variables we end up with are based entirely on the allocation of the rest for the donors that give aid to the recipient.

(education or health) aid to the sum of the prefix codes. 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 for this country (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 which is smaller than the sum of the prefix codes.[22] This seems to suggest data for Chinese Taipei contains a great deal of measurement error. Given that, in addition, data after 1996 is missing, this country in its entirety is dropped from the dataset. Yugoslavia, Sts Ex-Yugo. Unspec., 1996, has a negative education prefix sum. However, because total education aid is also negative, scaling should not be a problem for this observation. The same applies to Cayman Islands, 1991, in both sectors. For now, I keep these observations and simply apply the scaling; both observations 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. There are 4 observations (Costa Rica, 1992; Mexico, 1992; Panama, 1992; Saudi Arabia, 1991) for which the sum of the sectoral variables slightly exceeds DAC2a-RY disbursements (for some other observations this is because of rounding errors, occurring at the fifth digit after zero). This may arise when a recipient receives a negative total rest from a donor for which no sectoral allocation can be calculated. Since DAC2a-RY disbursements incorporate this negative amount of aid while the sectoral variables do not, the sectoral sum may exceed DAC2a-RY disbursements, if this negative value is not offset by positive amounts of rest from other donors for which the sectoral allocation is absent.

There is also one observation with a negative sectoral sum (Cayman islands, 1991). The reason is that for this observation the only rest variable that can be allocated across sectors is negative, whereas for the two donors with a positive rest value no sectoral allocation is available. Hence, each sectoral aid variable, and their sum, is negative, whereas the DAC2a-RY disbursements total is positive. I set all variables to missing for that observation.

Because the sum of the sectoral variables sometimes does not equal DAC2a-RY disbursements, as discussed above, the last thing I want to do in the data construction is scale the sectoral variables so that their sum matches aggregate disburse-

[22] The latter is also the case for Yugoslavia, Sts Ex-Yugo. Unspec., 1996, and Somalia, 1997.

ments received by a recipient in a given year. However, I cannot simply perform this scaling based on DAC2a-RY disbursements, which – recall – is derived by summing across donors in the recipient-donor-year dataset (DAC2a-RDY). This is because in the latter dataset only donors that are also available in DAC5 have been selected. Hence, DAC2a-RY disbursements may underestimate the total amount of aid a recipient receives.

Hence, I download grants and loans from DAC2a in a recipient-year format, selecting "all donors (total)" in the donor dimension. Missing grants are set to zero when loans are observed, and vice versa. Total disbursements are calculated as the sum of grants and loans (I call this data DAC2a-RY-AD, where AD stands for all donors). The correlation between this measure and DAC2a-RY disbursements (from the collapsed data) is extremely high (0.990). The sum of the sectoral variables has a similar high correlation with both these measures.

I scale the sectoral variables so that their sum equals the maximum of the sectoral sum, DAC2a-RY-AD disbursements and DAC2a-RY disbursements. Again, this follows the rationale that it is unlikely non-existing aid is reported, so the highest figure is likely to be the most accurate one. Obviously, education and health prefix codes are scaled along with total education and health aid. As already mentioned, there are four recipient-year observations (Costa Rica, 1992; Mexico, 1992; Panama, 1992; Saudi Arabia, 1991) that have negative rests in the recipient-donor-year dataset for at least one donor, which are not allocated across sectors because of a lack of information, and where this is not offset by unallocated positive rests. In such cases, the sectoral sum in the collapsed dataset (in recipient-year format) exceeds DAC2a-RY disbursements and may also exceed DAC2a-RY-AD disbursements in the recipient-year dataset (the latter is only the case for Panama, 1992). So, for these observations, I scale to the maximum of DAC2a-RY and DAC2a-RY-AD disbursements, as the sum of the sectoral variables is likely to exaggerate aid disbursements since 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).

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

Lastly, for Yugoslavia, Sts Ex-Yugo. Unspec., 1996, I turn all variables to missing. For this observation, some sectoral variables are negative while others are positive, roughly keeping each other in balance so that the sectoral sum is close to zero and very small in comparison to the DAC aggregate disbursements. As a result, the sectoral variables are scaled up to very large absolute values. The value for education aid, for instance, becomes -1204 million US $. Hence, all variables are set to missing for this observation.

Table 8 shows the scaling that takes place in the last step (calculated as the ratio of the sum of the sectoral aid variables to the aggregate disbursements variable that has been scaled to, as described above). This is compared to the scaling that would have taken place if we would have followed the practice of most previous papers that scale up the existing CRS data so that it matches aggregate DAC2a disbursements (scaling CRS is calculated with the sum of CRS sectoral variables in the numerator and the aggregate disbursements that have been scaled to in the denominator). As can be seen from the table, the difference between the two variables is large. This reflects the information added to the sectoral CRS data by the data construction described in this appendix. For the majority of observations, the scaling I perform in the last step of the data construction is limited in magnitude, and at the very least a lot smaller than if the CRS sectoral variables would have simply been scaled up, as is the case in most previous studies. This makes it more likely that the sectoral allocation of the data before scaling in the last step is a reasonable reflection of the actual sectoral allocation we would find if data from all donors were complete, making the scaling more defensible.

Table 8: **Scaling variables**

	Scaling	**Scaling CRS**
Mean	.7649	.3017
Observations	2390	2390
Standard deviation	.2063	.2668
Minimum	.0122	0
1st percentile	.0962	0
5th percentile	.3584	0
10th percentile	.4869	.0039
25th percentile	.6529	.0745
Median	.8079	.236
75th percentile	.9315	.4728
90th percentile	.9869	.7115
95th percentile	.9973	.8407
99th percentile	1	.9853
Maximum	1.128	1