THE IMPACT OF AID ON TOTAL GOVERNMENT EXPENDITURES

ISBN 978 90 361 0427 2

Cover design: Crasborn Graphic Designers bno, Valkenburg a.d. Geul

This book is no. **608** of the Tinbergen Institute Research Series, established through cooperation between Thela Thesis and the Tinbergen Institute. A list of books which already appeared in the series can be found in the back.

VRIJE UNIVERSITEIT

THE IMPACT OF AID ON TOTAL GOVERNMENT EXPENDITURES

ACADEMISCH PROEFSCHRIFT

ter verkrijging van de graad Doctor aan
de Vrije Universiteit Amsterdam,
op gezag van de rector magnificus
prof.dr. F.A. van der Duyn Schouten,
in het openbaar te verdedigen
ten overstaan van de promotiecommissie
van de Faculteit der Economische Wetenschappen en Bedrijfskunde
op vrijdag 17 april 2015 om 9.45 uur
in de aula van de universiteit,
De Boelelaan 1105

door Łukasz Marek Marć

geboren te Namysłów, Polen

promotoren:	prof.dr. J.W. Gunning
	prof.dr. C.T.M. Elbers

Acknowledgements

Many of my fellow PhD students and other colleagues are familiar with the PhD Comics (www.phdcomics.com). There is also a short story about PhD students in one of The Simpsons episodes when Marge Simpson says to her son: "Bart, don't make fun of grad students. They just made a terrible life choice".[1] While I find both of them funny, I feel that these stories are only vaguely related to my PhD time. And I am sure I did not make a terrible life choice. To a large extent it is thanks to the people I could count on during these five years in the Netherlands. I would like to express my gratitude towards them.

I would like to start by thanking my advisors, Chris Elbers and Jan Willem Gunning. It has been a long journey, but I could always count on their guidance and support. Jan Willem, thank you for your meticulous readings of my papers, for pushing me to connect them to a bigger picture, and for willingly sharing your profound knowledge. Chris, thank you for your great insights and ideas, for keeping me motivated, and for always having your doors open.

I have also received a fair deal of other good advice from my advisors (unlike the students in PhD Comics), especially during our department lunches in the “lovely”[2] VU canteen. Discussing my favorite TV show with them, I made a remark that the only disadvantage of watching The Wire is that some frequently used swear words come faster to my mind. Jan Willem commented: “Łukasz, keep them out of your papers.” I did. During my fieldwork in Ilorin, I met university friends of Emmanuel Olisadebe, Nigerian-born football player who ended up playing 25 games and scoring 11 goals for the Polish national team. One of my colleagues questioned the truthfulness of their stories, but Chris commented: “I always say that a good story doesn't need to be true to be good.”

[1] Type “Simpsons grad students" in youtube.

[2] Ironic. Very.

While I haven't applied this maxim to my research, it may prove beneficial in other situations.

I would also like to express my gratitude to the members of my PhD committee, Channing Arndt, Stephan Klasen, Robert Lensink, Menno Pradhan, and Ruerd Ruben. Thank you for taking time to read my thesis and for providing valuable comments.

Thanks for making the VU a fun and inspiring environment to my colleagues and friends from the Development Department: Melinda, Kimberley, Debra, Tahnee, Prita, Zlata, Lisette, Wendy, Berber, Martin, Remco, Menno, Jannis, Travers, and others, and to fellow PhD students and professors from the General Economics department, especially to Maarten for his regular visits that usually started with "You guys sit next to the printer so I will just bother you while waiting for my printout". Trudi, truly, you should be listed as my supervisor and advisor. I don't know how our group would function without you. And I greatly appreciated your daily visits. Thanks so much!

Wendy and Berber, thanks for inviting me to join the financial diaries project. It has been a great experience and a great complement to my me-and-my-computer research on the impact of aid. The times I spent working on this project in Nigeria and Kenya are definitely among the highlights of my PhD. Marijn and David, and the AIID team - it was great to work with you on this project. Greetings also to the Nigerian team. The funding for my fieldwork in Kenya by the C. Willems Stichting is greatly appreciated.

I am also grateful to Tinbergen Institute and its employees for creating a great studying environment, and to Dutch regulations regarding employment of PhD students. I could respond to my uncle's complaints about my prolonged studies by saying that I was actually full-time employed with a salary and pension contributions. "It's not a real job", he replied, nevertheless.

I spent the spring of 2014 visiting the United Nations University WIDER in Helsinki. *Kiitos* to the WIDER team for making it such a nice work place and immediately integrating all, even short term, visitors. Special thanks to Jukka for feedback and help, and to Heidi and Smriti.

The last three months of writing were intense and sometimes (not often, fortunately) stressful. I received a great deal of help in proof-reading and academic writing, valuable comments, and a lot of support, from Lisette, Zlata, Mark, Paul, and Marcin. Thank you, I wouldn't have been able to finish on time without you. Arturas, thanks for your econometric expertise. I am also grateful to Nadine, who was courageous enough to learn new words in Dutch (*fungibiliteit*) and who translated the summary into Dutch. "Your Dutch is brilliant" said my supervisor to me after reading it. Moreover, what kept me sane in these months were regular basketball games with Lithuanian friends. I wish we started playing earlier.

Above all, I would like to thank the friends that I met during my time in Amsterdam. I was lucky to share the 7A-42 office with two amazing fellow TIers and I feel we had a perfect mix of talking and working. Lisette, thank you for your kindness, support, great overseas travels (Petra!), and foremost for being a great friend. Paul, it was great sharing the office with you for the last three years. And your progress in Polish was amazing![3] Other friends I spent most time with are mostly also from the VU. Gosia, thanks for regular chats in Polish and introducing me to TI in 2009. Nadine, for your comments on reality, sharing important news, and helping me to understand the Netherlands better. Mark, for the regular off-the-track chats (also in the famous train to Rotterdam that we "missed" so badly afterwards), for our renting arrangement when our traveling schedules were perfectly complementary, and for not taking yourself seriously.[4] Zlata, your words of wisdom will always stay with me! Special thanks also to Lucy, Marcin, Stefi, Sergej, and Vili for great conversations and shared moments. TI was a fun place to study thanks to many people from the TI cohorts starting in 2008, 2009, and 2010. Thanks to Anita, David, Tomasz, and many many others.

Ogromne dziękuję bardzo to my friends back in Poland and in other parts of the world. According to my online flight diary, my most frequent travel routes have been: Eindhoven-Warszawa (12 flights), Eindhoven-Wrocław (12), Amsterdam-Warszawa (8). And this does not take into account about 10 *Jan*

[3]However, if I am ever obese I will blame you and your willingness to share M&Ms. When we were working temporarily in different offices, I kept receiving messages:"M&Ms have arrived, come!".

[4]Given that, let me tell you that I was warned about renting my place to Mark. Now, I must say that these worries were widely exaggerated. Fortunately, I was not liable for the glass doors he destroyed the first time he stayed at my place. One smashed flower and one call from my neighbor when I was in Kenya saying: "Łukasz! Your doors are wide open! There is noone inside!" are a low cost for helping a good friend and saving a few months of rent.

Kiepura train rides, and 3 miserable bus and car trips. Yes, I have visited Warszawa often. These trips were always a great way to clear my mind and charge my batteries. I have always had the feeling that there is a place I can go back to and immediately feel at home. Moreover, somehow I always felt more motivated to work afterwards. Andrzej and Kreska, thanks for thousands of talks, travels, and meetings. Monika, for sharing the important moments in life. Mateusz, for all these free-flowing never-ending conversations and a lot of inspiration and support. Robert and Svetlana, thank you for already 14 years of friendship even though the last time we lived in the same city was 2004. Gagatku, thanks for making me smile each time I see you or your photo and to M&T for making me part of your life. And there are so many more people to thank. I am very afraid I will miss someone, and since writing "thank you" shouldn't be stressful, then forgive me for grouping you (and allow me to thank you later in person). Thanks to my friends from TRAMP, for getting me bitten by the travel bug (apparently, that's how you translate: *zaszczepić bakcyla podróży*). To the SKN Ekonomii team for the passion for economics, great seminars, and friendships. To the ex-IBS team for making Imperium a fun place to work. To the SKN GEBR for regular meetings. And to those of you who I met regularly in Warszawa, somewhere in Poland, or abroad, for many great moments.

Last, but definitely not least, to my parents and family. Podziękowania dla Rodziców, Małego, Wujka Janusza i rodziny. Tradycja nakazuje, aby zacząć podziękowania od promotorów. Ale wiadomo, że w tych podziękowaniach ostatni są pierwszymi. Dziękuję za nieustające wsparcie, pomoc, życzliwość i zrozumienie. I za danie mi swobody do podejmowania własnych decyzji od kiedy pamiętam, także tych edukacyjnych.

Thank you! Dank jullie wel! Dziękuję!

Contents

List of Tables

CHAPTER 1

Introduction

1.1 Background

Foreign aid remains one of the most debated and the most controversial topics in development economics. While there seems to be a consensus that aid often has beneficial effects at a micro level, there is disagreement on aid effectiveness at the macro level. Therefore, there is still much debate on the aid-growth link. Most aid goes through the public sector. As a result, the channel that affects or even determines the impact of aid on the recipient's economy is likely the behavioral response of the recipient country's government to aid. To the extent that aid goes through the budget, aid can influence development outcomes (i) through increased spending (the quantity effect), or (ii) through improvement in the use of government resources because of donor involvement, or (iii) through lower taxes and borrowing (when aid substitutes other sources of revenues). Therefore, it is crucial to understand the impact of development assistance on the recipient country's government budget.

Two strands of literature investigate the behavior of the recipient country's government in response to the inflow of aid: on the one hand, the fungibility literature looks at aid and the allocation of public spending, and on the other hand there is the fiscal response literature, which looks at a broader fiscal relationship (spending and revenues). This PhD thesis contributes to both strands. I investigate the impact of aid on government expenditures and revenues, and estimate to what extent additional aid is fungible at the aggregate country level. If aid is fungible at the aggregate level, it raises government expenditures by less than the full amount.[5] This will happen when the recipient country's government adjusts its domestic revenues or borrowing, or if aid bypasses the budget. Fungibility may reflect differences in preferences between donors and recipients. The recipient government may disagree with the donor on the desirability of an increase in government expenditures, instead preferring to raise its citizens' private income by lowering taxes or reducing the government's borrowing needs. However, it may also happen that the recipient country's government is allowed to treat aid as fungible. In both cases, if on-budget aid is fungible at the aggregate level, at the margin it finances private consumption and savings, or reduces net borrowing.

Fungibility is more likely when monitoring the actual disbursement of aid is costly (Chatterjee et al., 2007). As Feyzioglu et al. (1998) suggest, it may seem that avoiding fungibility should be easy from a donor's perspective. Donors usually impose strict and carefully chosen conditions and a large part of aid is earmarked. Additionally, pre-intervention levels of government spending in aid-recipient countries are usually available to donors and could be used as an imperfect measure of what would happen if a particular aid package is not given. However, there are cases when this kind of monitoring is difficult. Usually countries receive aid from multiple sources and donor coordination is poor. At the same time, in developing nations, domestic resources fluctuate significantly within short periods of time and therefore past spending (and especially its composition) may not be informative of current or future spending.

Fungibility of aid is not necessarily bad for development. At the aggregate level, fungibility may limit diminishing returns of government spending that crowds out private investment. Furthermore, if aid leads to lower taxation, this may have a very high rate of return (at least in the short run) because in most developing countries distortions caused by taxes are very high (Devarajan et al., 2007; Feyzioglu et al., 1998).[6] At the sectoral, regional, or micro level, funds

[5] Aid is fully fungible if the government expenditures do not increase, and non-fungible if they increase by the full amount. Intuitively, I refer to the intermediate situation as: high fungibility, partial fungibility, and small fungibility.

[6] Conversely, it may also have negative macroeconomic consequences since aid is more volatile and unpredictable than domestic revenues. Increasing aid dependence reduces

may be reallocated to areas where investments are more productive. Moreover, there is a potential value for a country to make its own spending decisions.

What are the negative consequences of fungibility? From the donors perspective, fungibility means that aid is not spent as was intended[7] and it diminishes donor's level of control. Moreover, aid finances projects that would have been undertaken anyway, raising questions about money being wasted on earmarking. To overcome the unsatisfactory outcomes of conditionality, Collier et al. (1997) propose to allocate aid based on a retrospective assessment of major development outcomes (like growth). Devarajan and Swaroop (2000) suggest, rather than directly financing chosen sectors or regions, to use direct budgetary support conditional on an agreement about the composition and quality of government spending. From the researchers' perspective, when fungibility is present, assessing the effects of aid is more difficult.[8] At the macro level, aid's impact is the sum of the impact of increased spending and the impact of increased private consumption and savings (due to a tax decrease or corruption). The effect of aid is likely to be underestimated if a researcher ignores changes in private income or in increased public spending in a form not intended by the donor. These arguments suggest that measuring and understanding fungibility is important for donor strategies and may lead to more cost-efficient ways of aid distribution.

The most debated questions in the literature are (i) whether aid that is earmarked for one sector also finances expenditures in other sectors (categorical fungibility) and (ii) whether aid intended for development expenditures (public investment) ends up financing government consumption (general fungibility). Pack and Pack (1990, 1993) seek to examine these questions in Indonesia and the Dominican Republic, and reach different conclusions in their two papers. In one study, they find that foreign aid does not replace development expenditures in Indonesia. Instead, it even stimulates total government spending and own revenue. Moreover, categorical aid is spent as intended by

incentives for governments to introduce good policies and efficient institutions. Moreover, tax exemptions may be granted only to interest groups (Clements et al., 2004).

[7] Unless the donor gives aid as an unconditional budget support and allows the recipient country to use it as a substitute for domestic revenues or borrowing.

[8] Wagstaff (2011) discusses the challenges researchers face using an example of two World Bank health interventions that operated in Vietnam. The evaluation was difficult, because in response Vietnam's government intervened in provinces that the World Bank did not target.

donors (Pack and Pack, 1990).[9] However, Pack and Pack (1993) show that in the Dominican Republic, aid intended for development expenditures is used to finance deficit reduction, debt service, and taxes, implying high general fungibility. For India, Swaroop et al. (2000) report that, at the margin, aid is financing projects that would have been undertaken anyway, and that the freed resources are spent on non-development purposes. Van de Walle and Mu (2007) investigate whether aid intended to finance road rehabilitation in one Vietnam region also finances other regions at the margin. They find that the treatment communes built significantly more roads than the comparison group and that aid is only partly fungible between regions. Moreover, aid stuck to the road sector, so there was no sector fungibility.

In cross-country studies, Feyzioglu et al. (1998) find that concessional loans going to agriculture, education and energy are fungible, while those going to the transport and communication sectors are spent according to the donor's wishes. Devarajan et al. (2007) also show that aid going to the transport, communication, and energy sector is non-fungible. For the agriculture, health, and industry sectors, the hypothesis of full fungibility cannot be rejected. Pettersson (2007a) reports that over 65% of sectoral aid is fungible, as it is used outside the targeted sector. Conversely, Van de Sijpe (2013b) shows that off-budget health and educational aid (approximated by technical cooperation) is non-fungible as it has a very small impact on government expenditures in health and education sectors, while results for on-budget aid (approximated by a sector programme) are mixed. In a review of fungibility literature, Morrissey (2014) suggests that the extent of (general and categorical) fungibility is overstated because part of the aid never reaches the recipient country's budget.

Nevertheless, as discussed in more detail in chapter two, sector or general fungibility does not necessarily lead to fungibility at the aggregate level. Even if aid finances other sectors than intended by the donor at the margin, total government expenditures may increase by the full amount of aid. Morrissey (2014) points out that there is surprisingly little evidence on the additionality of aid, as fungibility studies focus more on whether aid is spent as intended by the donor, and fiscal response studies focus on the impact of aid on tax

[9] The case of Indonesia is special, because the government prepares a "Blue Book", which lists development project proposals that donors are invited to finance (Bastiaens, 2007).

revenues and borrowing. This PhD thesis attempts to fill this gap and provides new evidence on the impact of aid on government expenditures.

1.2 Thesis structure

This dissertation consists of three articles that aim to improve the understanding of the impact of aid on government expenditures. All chapters address the same main question, but use various econometric methods to investigate fungibility at the aggregate level. Moreover, this thesis contributes to the fiscal response literature by estimating the impact of aid on government revenues in chapter two and studing the causal links between government expenditures and aid in chapter four. In all three chapters, I use a rich recent dataset that I compiled using data from the OECD, the World Bank, the IMF, and the U.S. Census Bureau.

Inevitably, there is some overlap between the chapters, especially in the introductions. However, I leave the repetitive parts so that each chapter can be read on a stand-alone basis. Readers not familiar with the topic of fungibility at the aggregate level are advised to start with chapter two, which is the most comprehensive.

In chapter two, I investigate the short- and long-term impact of development assistance on government expenditures using a large recent dataset of 118 aid-recipients in the period 1980–2012. I begin by re-examining the fungibility at the aggregate level of total aid. If total aid is increasing government expenditures by less than its amount, it follows from the government budget constraint that aid substitutes domestic revenues, or that it substitutes borrowing, or that aid bypasses the budget. I test two out of these three channels. First, I provide additional evidence on the impact of aid on government revenues. Secondly, I distinguish between off- and on-budget aid and estimate the impact of both components on government expenditures. Thirdly, I disaggregate aid into bilateral and multilateral components and investigate whether their impact on government expenditures differs. To investigate the impact of aid or its components on government expenditures and revenues, I use the fixed effects model and the dynamic fixed effects model, hence I rely on non-instrumental methods of causal inference.

In chapter three I focus on two instrumental variable methods that have been used to account for potential endogeneity of aid in regressions estimating the impact of aid on government expenditures. First, I discuss the difference and system GMM estimators and employ them to estimate fungibility at the aggregate level using the same dataset as in chapter two. Moreover, I replicate and test the robustness of the results of Chatterjee et al. (2012). Secondly, I employ the other approach used in the fungibility literature and construct an instrument set based on geographical and cultural proximity between the donor and the recipient. I test the suitability of this method, again using the rich and recent dataset from chapter two.

In the fourth chapter, the causal links between on-budget aid and government expenditures are investigated on a dataset of 53 countries in the period 1980–2009. I test both links: the causal relation from aid to government expenditures, but also the impulses going from government expenditures to aid. The level of government expenditures can affect aid if donors react to budgetary situation in the recipient country or if the recipient country's government can decide on the timing of aid disbursement. In this chapter, I use Granger-causality methodology for panel data.

CHAPTER 2

The Impact of Aid on Government Expenditures: New Evidence on Fungibility

Aid is said to be fungible at the aggregate level if it raises government expenditures by less than the total amount. This happens when the recipient government decreases domestic revenue, decreases net borrowing, or when aid bypasses the budget. This study makes three contributions to both fungibility and fiscal response literature. Firstly, fungibility at the aggregate level is re-examined on a bigger recent panel dataset, distinguishing between short- and long-term impact of aid. The results indicate that aid is partly fungible in the long run and highly fungible in the short run. Secondly, to account for aid bypassing the budget, technical cooperation is used as a proxy for off-budget aid. Off-budget aid is found to be non-fungible and on-budget aid is partly fungible. Thirdly, fungibility of bilateral and multilateral aid is analyzed: the results indicate lower fungibility of multilateral aid.

I am especially grateful to Chris Elbers and Jan Willem Gunning for many useful comments and insights during numerous talks. I am also grateful to Maarten Lindenboom, Nicolas van de Sijpe, Sweder van Wijnbergen, anonymous referees, and seminar participants at Tinbergen Institute, VU University Amsterdam, Warsaw Economic Seminar, RGS conference, CSAE conference, and African Economic Conference for useful comments at the early stage of this research, and to Lisette Swart, Melinda Vigh, Jannis Visser, Lisa Winkler, Marcin Zamojski,

2.1 Introduction

Over the last 40 years aid[10] has accounted for a substantial part of developing countries' GDP and has been one of the main sources of government revenues in these countries. The empirical literature has tried (especially from the mid 1990s) to quantify aid's impact on growth and well-being in developing countries. The results were mixed in the studies up to 2009 (for reviews and discussion see, for example, Tarp 2006; Roodman 2007; Arndt et al. 2010). However, the most recent literature (see Arndt et al. 2010, 2011, 2014; Alvi and Senbeta 2012; Juselius et al. 2014; Mekasha and Tarp 2013; Galiani et al. 2014) mostly reports a positive impact of aid, although smaller than expected in the 1970s and 1980s.[11] A large part of aid goes to the public sector. Therefore, to better understand the impact of aid on growth and poverty reduction, it is crucial to determine the impact of aid on the behavior of one of the major actors in the development process, i.e. the recipient government.

Two strands of literature focus on the impact of development assistance on the behavior of recipient governments: fungibility studies and fiscal response studies. This chapter contributes to both. Fungibility is a broad term that describes situations when recipients respond to aid by changing the way they use their own resources. Aid is called fungible if the provision of goods and services intended by the donor is not fully achieved, because aid was used for other purposes.[12] McGillivray and Morrissey (2004) and Morrissey (2014) distinguish three types of fungibility: general fungibility, sector (categorical)

and especially to Zlata Tanović, for other help. I am grateful to UNU-WIDER for allowing me to use WIDER's premises and network to finalize this chapter, and to Jukka Pirttilä, Channing Arndt, Miguel Niño-Zarazúa, and internal seminar participants for useful suggestions. This chapter is a thoroughly revised version of an earlier paper (Marć, 2012). The usual disclaimers apply.

[10]This chapter uses standard terminology from the aid effectiveness literature. Official Development Assistance (ODA) is called aid, foreign aid, or development assistance. Also, the terms total aid and aggregate aid are used interchangeably to describe aid that is not disaggregated.

[11]Nevertheless, the debate is still ongoing. For example, Nowak-Lehmann et al. (2012) (and Herzer et al. 2014) find insignificant impact of aid on per-capita income, but their approach is criticized by Lof et al. (2014a) and Lof et al. (2014b).

[12]This is not the original meaning of this word, but this is the way fungibility is used in the literature. Owen Bader discusses the use of fungibility in development context in his blog: http://www.owen.org/blog/3224 (accessed May 2014). To avoid confusion, the current chapter defines the way the term fungibility is used in the introduction.

fungibility, and the additionality of aid.[13] General fungibility describes a situation when aid intended to finance public investment is diverted to government consumption. Aid is fungible on a sector (categorical) level if it is intended for one sector, but at the margin finances expenditures in other sectors.

This study focuses on the third aspect of fungibility: the additionality of aid.[14] It investigates by how much an additional euro of aid increases total government expenditures. At the aggregate level,[15] aid is fungible when one additional euro of aid increases total government expenditures by less than one euro, and fully fungible when government spending does not increase at all. This can happen when aid channeled through the budget substitutes for other sources of government revenue. This way the government is able to decrease taxes, decrease borrowing, or increase its reserves.

Morrissey (2014) points out that there is surprisingly little evidence on the additionality of aid. Although there is some evidence that aid is partly fungible at the aggregate level, the extent of fungibility of aid and its components and the importance of fungibility for the development process are still under debate. This study provides evidence on the extent of fungibility at the aggregate level, using a rich dataset of 118 countries for the period 1980–2012. I take into account the dynamic properties of the process, discuss endogeneity of aid in regressions, and explicitly distinguish between the short- and long-term impact of aid. The focus of the fungibility literature is on the short-term effects of aid: it tests whether aid increases government expenditures in the same year. However, aid may also have a long-term effect on the level of government expenditures. Investment in e.g. water supply and sanitation may reduce water-related diseases and in the long term decrease health

---

[13] A theoretical discussion of fungibility is presented in McGillivray and Morrissey (2000, 2001b).

[14] Note that the concerns that fungibility may be a red herring and that it distracts attention from more important issues, expressed by McGillivray and Morrissey (2000) and McGillivray and Morrissey (2001a), relate mostly to categorical fungibility. McGillivray and Morrissey (2001a) stress that "*the relevant concern is not fungibility* per se *but how aid impacts on total spending and financing and how spending plans are implemented*" (p. 118).

[15] Following Chatterjee et al. (2012) and Devarajan and Swaroop (2000) the term "fungibility at the aggregate level" is used interchangeably with "additionality of aid." Moreover, unless otherwise specified, fungibility refers to fungibility at the aggregate level, and not to general or categorical fungibility.

spending and total government expenditures. Moreover, since aid is usually committed for a longer period, donors may be more interested in a long-term impact of aid on government behavior rather than short-term adjustments. The inclusion of lagged government expenditures and data averaged over several years allow to estimate the long-term impact of aid on government expenditures. Furthermore, a change in government expenditures may be different depending on whether aid is increasing or decreasing, especially if the change in aid disbursement is unexpected. This study also tests the hypothesis of symmetric response to increases and decreases of aid.

If government expenditures do not increase by the amount of aid disbursed, it follows from the government budget constraint that domestic revenues adjust, or the net borrowing adjusts, or that aid bypasses the budget. This study investigates two out of these three channels: off-budget aid and government revenues.[16]

A large part of aid is never recorded in the budget, either because it is spent in the donor country, or because it directly reaches ultimate beneficiaries. The differences between these two components of aid should be taken into account, as on-budget aid has a direct impact on government expenditures, whereas the impact of off-budget aid is indirect because it is not recorded in the budget. Van de Sijpe (2013b) made the first attempt at distinguishing between recorded and unrecorded aid flows at the sectoral level. I follow this approach at the aggregate level and estimate the impact of on- and off-budget aid on government expenditures.

The impact of aid on government revenues is especially interesting for donors, as they are worried that aid may discourage tax effort and lead to lower domestic revenues of the government. This study provides additional evidence on the impact of aid on total government revenues and contributes to the fiscal response literature.

Fungibility of aid has been tested at the aggregate, sectoral, and regional level. Aid can also be disaggregated into bilateral and multilateral components. Ram (2003) points out that bilateral and multilateral aid differ for at least

[16]Since data on domestic revenues are not available, this chapter investigates the impact of aid on total government revenues. Moreover, reliable data on net borrowing are not available (see Appendix A).

three reasons: (i) donor motives, (ii) aid conditionalities, and (iii) closeness of the relationship between the donor and the recipient. Therefore, fungibility analysis should account for possible differences. To my knowledge, the fungibility of bilateral and multilateral aid has not been investigated using data from the last two decades. This study fills this gap.

The chapter proceeds as follows. The next section presents a literature review. Section three describes the data. Section four focuses on methodology and the choice of controls. Afterwards, the results for aggregate aid are presented in section five. Section six discusses and presents results for off- and on-budget aid, and section seven for bilateral and multilateral aid. The robustness of the results is tested in section eight. Section nine concludes.

2.2 Literature review

There is a large literature that investigates categorical and general fungibility.[17] Generally, aid is found to be fungible at both sector and general levels, however Morrissey (2014) points out that the extent of general and categorical fungibility is overstated because part of aid never reaches the recipient country's budget.

Fungibility detected at the sectoral or regional level does not mean that aid is fungible at the aggregate level. Aid intended for the health sector that is diverted to roads may still increase total government expenditures by the full amount. In this case aid is not fungible at the aggregate level. For example, Devarajan et al. (2007) report fungibility of aid at the sectoral level, while finding that at the aggregate level each dollar of aid leads to a 90 dollar cent increase in government expenditures in a sample of 18 Sub-Saharan African countries in the period 1971–1995. In a similar study, Feyzioglu et al. (1998) show that for their sample of 14 countries aid is not fungible at the aggregate level and that there is no tax relief that could be associated with fungibility. However, the number of observations is very small and, additionally, this result is not robust. Namely, when the number of countries is increased, aid appears

[17] See for example: Pack and Pack (1990, 1993); Feyzioglu et al. (1998); Swaroop et al. (2000); Van de Walle and Cratty (2005); Van de Walle and Mu (2007); Pettersson (2007a,b); Van de Sijpe (2013b,a). The literature reviews are provided by McGillivray and Morrissey (2001b); Morrissey (2014).

to be fungible.[18] Remmer (2004) uses an error correction model and finds a positive relationship between aid and government expenditures both in the short and in the long term. Arndt et al. (2014) investigate the cumulative effect of aid over 40 years treating the impact of aid on government consumption and government revenues as an intermediate result. They find a positive effect of aid on both measures.

Chatterjee et al. (2012) is the closest study to the current chapter. They analyze fungibility at the aggregate level and the impact of disaggregated aid (into investment aid, non-investment aid, and social infrastructure aid) on various measures of government expenditures. The major differences compared to the current study are: (i) their sample covers the period 1972–2000, and consists of 67 countries, (ii) off-budget aid is not taken into account, (iii) they do not distinguish between the short- and long-term impact of aid, (iv) I investigate bilateral and multilateral aid, they focus on investment and non-investment aid, and aid to social infrastructure. In addition to the fixed effects model, they use difference and system General Method of Moments (GMM) estimators, as well as IV methods (both discussed in chapter three). The dynamic specification is not used. They find that at the aggregate level the coefficient of aid's impact is 0.3, and the results are similar for the methods that try to account for the endogeneity of aid.[19]

Fungibility of aid at the aggregate level may indicate that recipient governments are decreasing their own tax efforts and that aid is substituting government revenues. As it is a particular concern for donors, parallel literature of fiscal response studies is investigating the impact of aid on government revenues and borrowing, often uses methods and datasets similar to the fungibility literature. Clements et al. (2004) find, for a sample of 107 countries in the period 1970–2000, that concessional loans increase domestic revenue,

[18] McGillivray and Morrissey (2000) point out concerns about this study.

[19] In the working paper version (Chatterjee et al., 2007), the link between fungibility and growth is made. Chatterjee and coauthors suggest that since there is no significant effect of foreign aid on growth in the presence of fungibility, fungibility can be the missing link that explains the mixed results of growth regressions. Conversely, Pettersson (2007b), using 57 country-specific estimates of fungibility incorporated into a growth model, finds no evidence that non-fungible aid is more effective for growth than fungible aid. Morrissey (2014) reviews the literature and notes that fungible aid does not seem to be less effective than non-fungible aid.

whereas grants are found to be fungible. Crivelli et al. (2012) confirm this result for a sample of 118 countries for the period 1980–2009 in a follow-up study. Clist and Morrissey (2011) change this specification slightly and conclude that both grants and loans are encouraging tax effort.

A wide review of the fiscal effects of aid (in particular, fungibility and fiscal response studies) was provided by McGillivray and Morrissey (2001b). Morrissey (2012, 2014) provides a recent literature review on the effects of aid on government spending and tax efforts. He points out that there is very little evidence on whether government expenditures increase fully by the amount of aid received, which is the main topic of this chapter.

The impact of bilateral and multilateral aid on government expenditures is analyzed by Cashel-Cordo and Craig (1990). They use a sample of 46 Least Developed Countries in the period 1975–1980. Bilateral aid, which is described as a relatively unconstrained type of development assitance, is found to be primarily a substitute for own government expenditures, while more constrained multilateral aid has a significant impact on expenditures. However, Gang and Khan (1990), analyzing India's expenditures, find that grants, loans, and multilateral aid have no significant impact on government expenditures, while bilateral aid induces transfers of domestic public resources from non-investment to investment for development purposes.

2.3 Data

This study employs the most recently available data that cover the period 1980–2012 for 118 countries. The analysis is limited to countries that have received aid inflows in this period and are listed as Least Developed Countries and Other Low Income Countries (together called LDCs henceforth), Lower and Middle Income Countries and Territories (LMICs), and Upper and Middle Income Countries (UMICs). Countries were classified according to the World Development Classification for 2009 (see Appendix B for the list of countries and a discussion). The 118 countries in the sample can be classified as follows: 35 LDCs, 50 LMICs, and 33 UMICs. There are 49 African countries, 31 are from Asia, 8 are from Europe, 7 are from Oceania, and 23 are from Central and South America. Compared to the previous studies that analyzed fungibility at the aggregate level presented in the literature review, the sample covers more countries and more years, which results in a substantially larger number of observations. Moreover, it can be expected that the data collection mechanisms have improved over the years, and thus that the recent data are of a better quality.

The data and sources are described in more detail in Appendix A. In general, the variables used in this study can be divided into three types: public spending variables, aid variables, and control variables. The dependent variable is total government spending or government revenue[20] as a share of GDP from the International Monetary Fund's (IMF's) World Economic Outlook (WEO). Among the independent variables, the Official Development Assistance (ODA) is of the main interest. The data for aggregate aid disbursements, as well as for bilateral aid, multilateral aid, and technical cooperation (all in current dollars) are obtained from OECD's Development Assistance Committee database, Table DAC2a.[21] All aid variables are expressed as a share of the aid-recipient's GDP, where GDP (in current dollars) is taken from the WEO. The control variables are obtained mostly from the World Development Indicators (WDI): agricultural value added as a share of the aid-recipient's GDP, annual growth of GDP (%), GDP per capita in constant dollars (PPP), literacy rate, annual population growth (%), population, and sum of exports and imports of goods and services as a share of GDP. Furthermore, the annual inflation rate (%) is obtained from the WEO. Finally, the U.S. Census Bureau's international database (IDB) provides the data on infant mortality rates.

The control and aid variables are almost complete for the whole period. Therefore, the sample size is determined by the availability of the dependent variables: for many countries government expenditures and government revenues are missing in the 1980s, and for some countries in the 1990s. The data are almost complete from 2000 onwards.

Table 1 reports the means of the main variables. The unweighted share of government expenditures in GDP is the smallest among LDCs and at a similar

[20]Government revenue is described as the sum of "taxes, social contributions, grants receivable, and other revenue" (WEO dataset). Therefore, it also includes at least part of on-budget aid. It does not include net borrowing and lending. To make the distinction clear, the term domestic revenues is used to describe government revenues minus any financing from abroad.

[21]Available at http://stats.oecd.org/, under Development → Other.

Table 1: The share of government expenditures and aid by several characteristics, unweighted (Un) and weighted by GDP (W)

	Gov. (% GDP)		Aid (% GDP)		Bil aid (% GDP)		Mul aid (% GDP)	
	Un	W	Un	W	Un	W	Un	W
All	28,49	24,26	8,58	0,90	5,46	0,58	2,89	0,29
LDCs	21,92	19,10	13,59	9,01	7,87	5,43	5,47	3,47
LMICs	31,82	22,80	8,69	0,59	6,12	0,41	2,29	0,17
UMICs	30,59	30,17	2,77	0,46	1,71	0,30	0,94	0,15
Africa	26,60	28,50	10,75	3,20	6,26	1,98	4,27	1,17
Asia	26,49	22,07	5,47	0,51	3,19	0,35	1,81	0,14
Europe	37,99	37,90	3,53	0,71	1,91	0,36	1,47	0,33
Oceania	45,75	32,45	21,94	8,31	18,48	7,24	3,38	1,05
Latin America	25,55	25,75	4,10	0,70	2,53	0,52	1,53	0,18
1980-1989	27,85	23,06	8,30	1,17	5,13	0,70	2,66	0,40
1990-1999	27,71	20,70	9,40	1,16	5,86	0,74	3,39	0,39
2000-2012	28,87	25,25	8,19	0,81	5,40	0,54	2,68	0,26
China	19,96	20,15	0,24	0,13	0,16	0,09	0,07	0,03
India	25,66	26,27	0,48	0,28	0,24	0,16	0,24	0,12

Notes: own calculations based on IMF and OECD data. Total aid includes aid from non-DAC (Development Assistance Committee) donors.

level among LMICs and UMICs. Europe and Oceania have a much higher level of expenditures than the remaining three continents. For the total sample the share of government expenditures remained relatively constant over the last 30 years. As expected, LDCs receive substantially more aid in relation to their GDP than higher income countries. Therefore, Africa and Oceania, continents with a large number of LDCs, are the biggest aid recipients in relation to their GDP.[22]

Unweighted averages do not account for the fact that the biggest (both in terms of population and the size of economy) countries receive less per capita aid, and aid accounts for a small share of GDP. For example, India's

[22] Due to the concerns about the data quality and due to the fact that off-budget aid accounts for a very small (less than 0.5%) share of GDP, non-DAC members donations are not included in the analysis of disaggregated aid.

share of aid surpassed 1% of GDP only in the early 1980s, while for China the highest share was in 1992 and equalled 0.6%. Table 1 also presents calculations weighted by total GDP (giving higher weight to bigger economies, but also to recent years, when GDP in most countries was higher than in the 1980s). When the size of the economy is taken into account, aid accounts for 0.90% of aid-recipients' GDP. The difference between LDCs and LMICs is even more stricking: aid constitutes more than 9.01% of GDP in the LDCs, which is more than 15 times the percentage it constitutes in the LMICs (which include China and India). In addition, differences between continents are substantial: Europe, South America, and Asia—continents with almost exclusively LMICs—have low shares of aid in GDP. Africa, having the majority of LDCs, and Oceania have a relatively high share. As a robustness check, I include regressions weighted by population and GDP to account for different patterns of the distribution of aid with respect to the size of countries.

2.4 Methodology and empirical framework

Analytical framework and the basic fixed effects model

The government budget constraint in any given period can be written as follows:[23]

$$GovernmentExpenditures = Revenues + Aid + NetBorrowing \quad (2.1)$$

where *Revenues* are domestic revenues from taxes, natural resources, and other sources. In this equation, *Aid* represents the part of the ODA that goes through and is recorded in the budget. It follows from this equation that in response to an increase of on-budget aid, government can either: (i) adjust its expenditures, (ii) adjust domestic revenues, (iii) adjust borrowing, or do any combination of the three. These possibilities are discussed by Clements et al. (2004).

At the aggregate level, the fungibility literature focuses on the impact of aid on total government expenditures. The approach most widely used in the

[23]The analytical framework follows e.g. Clements et al. (2004) with modifications.

literature is a panel data regression with country and year fixed effects, using annual data. For aggregate aid, the reduced form equation is the following:

$$GovExp_{it} = \alpha AID_{it} + \beta X_{it} + \gamma_t + \delta_i + \varepsilon_{it} \tag{2.2}$$

where $GovExp_{it}$ represents the share of government expenditures in GDP and AID_{it} the share of ODA (or its components) in GDP. X_{it} is a set of controls, γ_t are year fixed effects, δ_i represents the recipient country's fixed effects, and ε_{it} is an error term. Subscript 'i' refers to a recipient country and 't' to year. In this model, α is the coefficient of main interest: if α is equal to 1, aid is non-fungible. A value of α less than 1 but greater than 0 means that aid is partially fungible, and α equal to 0 indicates that aid is fully fungible.[24] In these cases, $(1-\alpha)$ gives the extent of fungibility. Finally, α bigger than 1 provides evidence for the "crowding-in" effect (often referred to as the "flypaper effect" in the fungibility literature),[25] the situation in which one euro of aid leads to more than one euro increase in government expenditures.[26]

It is important to analyze the implications from the reduced form equation (2.2) in light of equation (2.1). If aid is estimated to be fully fungible, so that an increase in aid does not lead to an increase in government expenditures, this means that domestic revenues or net borrowing (or both) decrease. However, if total aid is not disaggregated into on- and off-budget components in equation (2.2), an estimate of α smaller than one may also mean that aid bypasses the budget (off-budget aid is discussed in Section 2.6).[27]

[24] Formally, full fungibility occurs if the impact of aid is not greater than the marginal effect of unconditional resources (domestic revenues and net borrowing) (Devarajan et al., 2007). However, as neither data for the domestic revenues, nor for the net borrowing are available, I assume that the marginal effect of unconditional resources is zero.

[25] Originally, the "flypaper effect" was used to describe the empirical phenomenon that grants from the higher level government to local governments increase local spending by more than local income of equivalent size would (see e.g. Hines and Thaler 1995). In the fungibility literature, the flypaper effect is also quite often used to describe a situation when recipient's expenditures increase by more than the amount of aid received (Leiderer, 2012; Pettersson, 2007b). See McGillivray and Morrissey (2000) for discussion.

[26] McGillivray and Morrissey (2001a) offer four theoretical scenarios where the inflow of aid leads to a greater increase in public expenditures than the value of aid in recipient countries. These may happen because of misconceptions or illusions regarding either the real or nominal value of the aid inflow. They show that it is even possible if own government revenues decrease.

[27] It may also happen that the estimated impact of aid is downward biased because part of aid is embezzled before it reaches the government budget (and before it is recorded there).

Equation (2.2) is a starting point to investigate the impact of aid on government expenditures using the panel dataset described in Section 2.3. Bertrand et al. (2004) and Cameron and Trivedi (2005) point out that outcomes are serially correlated in panel data. When serial correlation in errors is not accounted for, standard errors are underestimated and the null hypothesis of no effect is rejected too often. I use the Wald test to check for serial correlation of errors in the linear panel data model. The null hypothesis of no correlation is equivalent to the residuals having -0.5 autocorrelation in the first difference model. For all specifications, the hypothesis of no serial correlation is rejected. This result is confirmed by the autocorrelation test designed by Arellano and Bond (1991). On the basis of Monte Carlo simulations, Bertrand et al. (2004) suggest using a heteroskedasticity and autocorrelation consistent (HAC) estimate for the variance-covariance matrix as a solution to this problem. This solution works well when the number of groups is sufficiently large, which can be reasonably assumed in this sample consisting of 118 countries. Therefore, to account for autocorrelation, standard errors are clustered at the country level.

Dynamic panel data model

Van de Sijpe (2013b) points out that the presence of serial correlation may indicate the dynamic misspecification of the model. A second indication that the model may be dynamically misspecified comes from the comparison of fixed effects (FE) estimates (within-estimator) and the first difference (FD) model. The first difference estimates are found to be smaller than the fixed effects estimates suggesting a violation of the strict exogeneity assumption.[28] This indicates that both FD and FE estimates are inconsistent with different probability limits.

The autocorrelation and differences between FD and FE estimates suggest that a model of public expenditures requires a dynamic specification. Indeed, the correlation of current and lagged values of government expenditures is very high. Inclusion of the lagged dependent variable may solve this problem and remove autocorrelation in errors.

[28] Detailed results are available upon request.

Roodman (2009a) points out that when the number of periods is large, dynamic panel bias becomes insignificant[29] and the fixed effects estimator can be used. Compared to GMM models discussed in chapter three, this estimator requires fewer assumptions and the results are less sensitive to the researcher's choices. Therefore, in addition to a standard fixed effects model presented in the previous section, a fixed effects model with a lagged dependent variable is used as well. This model is of the form:

$$GovExp_{it} = \rho GovExp_{it-1} + \alpha AID_{it} + \beta X_{it} + \gamma_t + \delta_i + \varepsilon_{it} \tag{2.3}$$

The inclusion of a lagged dependent variable not only solves the dynamic misspecification problem, but also allows to estimate a long-run effect of foreign aid on government expenditures. In the steady state, equation (2.3) can be rewritten as:

$$GovExp_i = \frac{\alpha}{1-\rho} AID_i + \frac{\beta}{1-\rho} X_i + \frac{\delta_i}{1-\rho} + \frac{\varepsilon_i}{1-\rho} \tag{2.4}$$

The presence of a lagged dependent variable changes the way fungibility coefficients should be analyzed. In this equation α measures the short-term adjustment effect of aid on government expenditures, whereas the long-run effect is equal to $\alpha_{LR} = \alpha/(1-\rho)$.

Alternatively, to estimate medium- and long-term impact of aid on government expenditures, data are averaged over three, four, and five years, after which equations (2.2) and (2.3) are estimated. This approach has been widely used in the parallel literature on the impact of aid on growth, since averaged data reduce both business cycle effects and measurement errors. Furthermore, Arndt et al. (2010, 2014) argue that the impact of aid on growth is a long-term process, so that it is more appropriate to use long panels, or even to use averages of time-series of all available data for each country.

[29]Dynamic panel bias is a result of correlation between a regressor and the error term caused by the demanding process of mean subtraction in the within estimator. Nickell (1981) discusses this bias. For a reasonably large value of T the bias approximates to $(\hat{\rho}-\rho) \simeq -\frac{1+\rho}{T-1}$, so the persistence of y is always underestimated for $\rho > 0$. For $T = 10$ and $\rho = 0.5$ the bias will be -0.167, and for T around 30 it will be three times smaller and close to -0.05. The fact that the lagged dependent variable is likely underestimated means that the long-term impact of aid derived in the following equations may be, *ceteris paribus*, also underestimated. However, the direction of the bias of aid and other parameters in the fixed effects model is ambiguous (Juodis, 2013).

For the relation between aid and government expenditures, averaging can also help to reduce the problem of measurement error and account for a business cycle that may influence both the level of aid and the level of government expenditures. What is more, if the actual level of aid disbursement in a given year is not known to the government before the budget is planned, the estimate of the immediate response of government expenditures to aid may reflect a government's short-term capacity to manage additional resources rather than government's preferences over the optimal aid allocation. Unexpected inflows and outflows may be averaged out when longer time periods are used. As a result, an estimate of the impact of averaged aid on averaged government expenditures will better reflect government preferences.

However, averaging comes at the cost of decreasing the number of periods in a panel data setting. The estimates from equation (2.3) may be biased due to the inclusion of a lagged dependent variable (the dynamic bias decreases with the number of periods, as previously noted).[30] The difference and system GMM estimators which are used to account for the dynamic panel bias are discussed and applied to the same dataset in chapter three.

Endogeneity

The process of aid determination makes it very likely for aid to be endogenous in various regressions. The issue of endogeneity has dominated the literature on the impact of aid on growth in the last 15 years. Various methods have been proposed to account for endogeneity, ranging from internal instruments in the difference or system GMM estimation (Alvi and Senbeta, 2012; Annen and Kosempel, 2009; Rajan and Subramanian, 2008), through a standard set of instruments that include recipient country characteristics and log of population size (proposed by Boone 1996), and often include also the donor-recipient interaction (among others: Burnside and Dollar 2000; Hansen and Tarp 2000, 2001; Dalgaard et al. 2004; Rajan and Subramanian 2008; Arndt et al. 2010, 2014), to quasi-experiments (Galiani et al., 2014). The majority of studies find evidence for the positive impact of aid on growth and argue that the new and improved instruments are valid. However, Clemens et al. (2012), reviewing

[30]This does not apply to the results from equation (2.2).

instrumental variables used in aid-growth studies, conclude that the search for strong instrument that does not raise important questions about its validity has not finished.[31]

Clearly, aid may also be endogenous when the impact of aid on total government expenditures is analyzed:

1. The level of government expenditures may be decided simultaneously with the level of aid inflow: McGillivray and Morrissey (2000) suggest that often the recipient can decide on the timing of aid disbursement. Then, e.g. a negative shock to revenues may cause increased inflows of aid.
2. Donors may focus on the provision of merit goods in countries that fail to provide them.
3. Countries with good policies may be allowed to treat part of their aid as fungible (Pettersson, 2007b).

However, in spite of these important reasons, it can be argued that the problem of endogeneity is less severe than in aid-growth studies. The simultaneity problem (point 1) can be reduced by averaging aid and government variables over time. Fixed effects can account for aid dependency or the indirect effect of good policies (point 3), provided these do not change over time. Even if donors target countries that fail at providing merit goods (point 2), that does not necessarily mean that the targeting will affect the level of government expenditures and that donors target also total expenditures.

Thus, there is a potential bias due to endogeneity. However, because this bias is expected to be small and because of the previously mentioned concerns about instrument quality, this study relies on non-instrumental methods of causal inference only. The usefulness of two instrumental methods used in the fungibility literature (the difference/system GMM and the instrument based on cultural and geographical proximity) is discussed in chapter three.

Asymmetric response to increases and decreases

Estimating the impact of development assistance also raises an interesting empirical question that has been largely ignored in the fiscal response literature:

[31] Further improvements of instrumentation strategies have been suggested by Bazzi and Clemens (2013). For a criticism of instrumentation strategies in the aid-growth literature see Deaton (2010).

whether the response of the recipient government to increases and decreases in aid levels is asymmetric. The behavioral response of a government depends on various factors, including: i) government preferences, ii) aid conditionalities, iii) access to financial markets, iv) and access to other sources of revenue. For example, if the access to financial markets or other sources of revenue is limited, a decrease in aid levels, especially an unexpected one, may have a large and immediate impact on the level of government expenditures. Conversely, additional aid may simply be stored at the central bank, without increasing the government expenditures.

The asymmetric response to increases and decreases is a popular topic in the literature analyzing resource price changes (Bachmeier and Griffin, 2003; Kilian and Vigfusson, 2011; Radchenko, 2005) and intergovernmental grants (Gamkhar and Oates, 1996; Stine, 1994). Following the latter literature that uses fixed effects models, I use two specifications. Firstly, change in aid is included if aid increases:

$$GovExp_{it} = \alpha_1 AID_{it} + \alpha_2 I_t (AID_{it} - AID_{it-1}) + \beta X_{it} + \gamma_t + \delta_i + \varepsilon_{it} \quad (2.5)$$

$$I_t = 1 \quad if \quad AID_{it} > AID_{it-1}, \quad and \quad I_t = 0 \quad otherwise$$

where I_t is a dummy variable equal to 1 when the share of aid in GDP increases. The null hypothesis is $\alpha_2 = 0$, which indicates symmetric response to increases and decreases. If the null hypothesis is rejected, α_1 is the response of government expenditures to an aid decrease, and α_2 is the response of government expenditures to an increase in aid, whereas α_1 to an increase in aid level.

Secondly, change in aid is included if aid decreases:

$$GovExp_{it} = \alpha_1 AID_{it} + \alpha_2 D_t (AID_{it} - AID_{it-1}) + \beta X_{it} + \gamma_t + \delta_i + \varepsilon_{it} \quad (2.6)$$

$$D_t = 1 \quad if \quad AID_{it} < AID_{it-1}, \quad and \quad D_t = 0 \quad otherwise$$

where D_t is a dummy variable equal to 1 when the share of aid in GDP decreases. Again, the null hypothesis is $\alpha_2 = 0$, indicating symmetric response to increases and decreases. If the null hypothesis is not rejected, α_1 is the response of government expenditures to aid increase. If the null hypothesis is rejected, α_1 measures the response to an aid level decrease, and α_2 to a reduction in aid.

Alternatively, the difference in the response can be estimated directly from:

$$GovExp_{it} = \alpha_1 I_t AID_{it} + \alpha_2 D_t AID_{it} + \beta X_{it} + \gamma_t + \delta_i + \varepsilon_{it} \qquad (2.7)$$

where α_1 is the impact of aid increases on government expenditures, and α_2 is the impact of aid decreases.

Controls

Variables that may explain the size of government expenditures, and hence provide a more precise estimate of the coefficient of main interest, and/or variables that are correlated with aid and cause estimates to be biased if not controlled for, are included in the initial set of controls. Plausible controls are tested in various configurations, both for aggregated and disaggregated aid.[32]

Five variables are used as controls: annual GDP growth in percentages (GDP growth), lagged annual inflation rate (L.Inflation), lagged agricultural value added as a share of GDP (L.Agr VA), lagged sum of imports and exports of goods and services as a share of GDP (L.Exp+imp in GDP), and lagged infant mortality rate per 1000 births (L.Infant mortality).[33]

Following Feyzioglu et al. (1998), Clements et al. (2004), and Chatterjee et al. (2012), agricultural value added is used to control for spending in the agriculture sector and, additionally, for the level of development. The infant mortality rate serves as a proxy for health-care and social security spending.

The growth variable should control for the response of government spending to short-run shocks in GDP per capita (Van de Sijpe, 2013b). As has been noted before, McGillivray and Morrissey (2000) suggest that recipients tend to have a large freedom of choice over the extent to which committed aid is disbursed in a single year. Hence, any shock to expenditures (proxied by the growth rate) would also affect the amount of aid disbursed, and aid would be correlated with the error term.

[32]Plausible controls are discussed in Appendix C. Regression results including these controls are discussed in Section 2.8.

[33]Some control variables may be non-stationary, for example agricultural value added or the infant mortality rate. Im-Pesaran-Shin and Fisher tests – stationarity tests that allow unbalanced panels – reject the null hypothesis that all panels contain a unit root for all variables in the sample. Moreover, the impact of aid on government expenditures remains similar when all control variables are excluded.

A ratio of a sum of imports and exports to GDP is used to capture the effect of openness of the economy on government expenditures. Rodrik (1998) shows that the size of government has been larger in most open economies due to the risk of external shocks. Alesina and Wacziarg (1998) argue that this positive correlation is rather due to the country size, since trade openness and government expenditures are negatively correlated with the share of public consumption in GDP. This result was questioned recently by Ram (2009) who suggests that Rodrik's (1998) explanation may be the correct one. To show that, he uses a fixed-effects format to account for cross-country heterogeneity that was not taken into account by Alesina and Wacziarg (1998).

The inflation rate is included as a factor that has been found to be significantly related to cross-country variations in domestic revenues (Clements et al., 2004). It can increase revenues either when taxes are not indexed or through seigniorage.

Other controls include fixed effects that account for time-invariant factors like geography, colonial history, or the legal system. Additionally, fixed effects are expected to account for differences in reporting scopes of government expenditures (if the reporting methodology is constant over time). Year fixed effects are used to control for events that had an impact on all recipients in a given year.

Lagged control variables are used to account for potential simultaneity problems. The only exception is made for GDP growth, since it accounts for immediate responses of government expenditures to shocks. As a robustness check, current values of the control variables are used.

2.5 The impact of aggregate aid on government expenditures

The first part of the analysis focuses on the impact of total aid on government expenditures. As noted before, an estimated coefficient of the share of aid smaller than one means that aid is partly fungible (fully fungible when the coefficient equals zero). If fungibility is detected, it means that part of aid is financing tax reductions, decreasing net borrowing, or bypasses the budget. As can be seen from Table 2, in the basic fixed effects specification, which is similar to specifications used in other fungibility studies (column [2]), fungibility of total aid $(1-\alpha)$ is high for the sample of all countries and equals 0.855. This means that when the share of aid in GDP increases by one percentage point (pp), government expenditures as the share of GDP increase by 0.145 pp. The remaining 0.855 pp substitute expenditures, i.e. are used to decrease taxes or borrowing needs, or bypasses the budget. Compared to the pooled Ordinary Least Squares (OLS) model (column [1]), which does not include country fixed effects, the estimate is smaller.

However, the Arellano-Bond test of autocorrelation rejects the null hypothesis of no autocorrelation for both models presented in columns [1] and [2], which suggests misspecification. As expected, including the lagged dependent variable accounts for autocorrelation. Tests for both first- and second-order serial correlation fail to reject the null of no autocorrelation in all three specifications with the lagged dependent variable. As has been noted before, even when strict exogeneity is violated due to the inclusion of a lagged dependent variable, for a long panel the dynamic bias becomes insignificant and the fixed effects estimator can be used (Roodman, 2009a).[34]

The inclusion of the lagged dependent variable not only solves the problem of autocorrelation, but also allows to distinguish between short-term adjustments and long-term impact of aid on government expenditures. The fixed effects model in column [3] shows that approximately 92% of aid is fungible at the aggregate level in the short-run and 78% in the long run. This is confirmed in column [4] which includes total aid squared. Table 3 shows the results for three, four, and five year averages of data. These results suggest that the long-run estimate of the impact of aid on government expenditures is underestimated in the models that use annual data. The coefficient of short- (or medium-) term impact of aid increases in the number of years that the data are averaged over, and so does the estimate of the long-term impact. These results suggest that in the long run aid is partially fungible and government expenditures increase by 40–50% of aid.

Fungibility may be a non-linear function of aid. Pack and Pack (1993) suggest that fungibility may be inversely related to the share of aid in government

[34]Alternatively, the difference and system GMM estimators can be employed. See chapter three for the discussion.

Table 2: The impact of aid on government expenditures and revenues

	[1] EXP (OLS)	[2] EXP (FE)	[3] EXP (FE)	[4] EXP (FE)	[5] REV (FE)	[6] REV (FE)
Total aid	0.517***	0.145***	0.081***	0.084*	0.061	0.079**
	(0.19)	(0.05)	(0.02)	(0.05)	(0.06)	(0.03)
L.Gov. exp. in GDP			0.622***	0.621***		
			(0.03)	(0.03)		
GDP growth	-0.102	-0.035	-0.006	-0.006	0.115**	0.100**
	(0.09)	(0.04)	(0.03)	(0.03)	(0.06)	(0.04)
L.Inflation	0.004	0.002	0.000	0.000	0.000	-0.001
	(0.01)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
L.Infant mortality	-0.060*	0.013	-0.001	-0.001	-0.006	-0.004
	(0.03)	(0.03)	(0.01)	(0.01)	(0.03)	(0.02)
L.Exp+imp in GDP	0.037*	0.008	0.005	0.005	0.002	-0.003
	(0.02)	(0.02)	(0.01)	(0.01)	(0.02)	(0.01)
L.Agr. VA	-0.274***	-0.209***	-0.040	-0.040	-0.207***	-0.068
	(0.06)	(0.07)	(0.03)	(0.03)	(0.08)	(0.05)
Total aid sq.				-0.363		
				(4.48)		
L.Gov. rev. in GDP						0.538***
						(0.09)
Constant	39.458***	27.167***	7.698***	7.690***	25.137***	11.817***
	(6.41)	(4.68)	(2.85)	(2.87)	(5.29)	(3.57)
LR total aid			0.216***	0.221*		
			(0.07)	(0.14)		
R-sqr	0.27	0.14	0.48	0.48	0.15	0.35
F-statistic	8.07	7.60	62.91	62.41	5.02	31.20
Obs.	2112	2112	2024	2024	2132	2045

Notes: The dependent variable is the share of government expenditures in GDP (%) (denoted as EXP) or the share of general government revenue in GDP (%) (denoted as REV). *Total aid* is measured as the share of aid in GDP (%). Year fixed effects are included in all models (also OLS), country fixed effects are included in the fixed effects model (FE), coefficients are not reported. Standard errors are clustered by country and are reported in parentheses. LR presents the long-run impact of aid on government expenditures from equation (2.4). $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$.

expenditures since large aid flows decrease monitoring costs. To test this non-linearity, total aid squared is included as a regressor in column [4] of Table 2. The coefficient of aid squared is negative, although not significantly different from zero. The same results are found for OLS and FE models without the lagged dependent variable, as well as for the averaged data (not reported). Only in some specifications the coefficient is significantly different from zero. Thus,

Table 3: The impact of aid on government expenditures - averaged data

	[1] 3y	[2] 3y	[3] 4y	[4] 4y	[5] 5y	[6] 5y
Total aid	0.241***	0.206***	0.283***	0.265***	0.363***	0.402***
	(0.08)	(0.05)	(0.08)	(0.07)	(0.11)	(0.12)
L.Gov. exp. in GDP		0.378***		0.311***		0.188**
		(0.06)		(0.08)		(0.08)
GDP growth	-0.090	-0.019	-0.061	-0.035	-0.052	-0.019
	(0.07)	(0.08)	(0.08)	(0.08)	(0.08)	(0.11)
L.Inflation	0.000	0.002	-0.000	-0.001***	-0.001*	-0.001
	(0.00)	(0.01)	(0.00)	(0.00)	(0.00)	(0.00)
L.Infant mortality	0.005	-0.002	-0.007	-0.014	-0.032	-0.037
	(0.03)	(0.02)	(0.03)	(0.03)	(0.03)	(0.04)
L.Exp+imp in GDP	-0.006	-0.001	-0.003	-0.001	-0.019	-0.004
	(0.02)	(0.02)	(0.02)	(0.03)	(0.02)	(0.03)
L.Agr. VA	-0.219***	-0.138**	-0.213***	-0.176**	-0.230***	-0.169*
	(0.08)	(0.06)	(0.07)	(0.09)	(0.07)	(0.09)
Constant	33.310***	20.639***	33.045***	21.692***	33.345***	27.652***
	(3.23)	(3.39)	(3.36)	(4.96)	(4.04)	(3.42)
LR total aid		0.331***		0.384***		0.494***
		(0.09)		(0.10)		(0.14)
R-sqr	0.18	0.33	0.21	0.32	0.25	0.32
F-statistic	9.33	17.96	12.09	29.71	14.64	14.02
Obs.	734	650	550	474	434	364

Notes: The dependent variable is the share of government expenditures in GDP (%). *Total aid* is measured as the share of aid in GDP (%). Year and country fixed effects included in all models, coefficients are not reported. Standard errors are clustered by country and are reported in parentheses. LR presents the long-run impact of aid on government expenditures from equation (2.4). 3y, 4y, 5y means data averaged over 3, 4 and 5 years respectively. $^{*}p < 0.1, ^{**}p < 0.05, ^{***}p < 0.01$.

in contrast to what Pack and Pack (1993) suggest, evidence of non-linearity is not found in this analysis. However, the negative coefficient suggests that fungibility can be higher for high amounts of aid: the more aid is given to a particular country, the lower is the impact of aid on the level of government expenditures.

The impact of aid on government expenditures may differ depending on whether aid is increasing or decreasing. Section 2.4 discusses potential reasons for the asymmetry in response and offers three ways to investigate this hypothesis. Table 11 in Appendix D reports the results. At a 5% level the null hypothesis of symmetric response cannot be rejected (see columns [1]–[4]). Also, the estimates of the impact of aid when it is decreasing are close to the estimates when aid is increasing. Therefore, there is no evidence for the hypothesis of an asymmetric response. This means that when the share of aid in GDP decreases, the recipient country's government is able to increase domestic revenues or net borrowing to partly offset the decrease in aid.

As pointed out in Section 2.4, fungibility of aid at the aggregate level means that own government revenues adjust, or net borrowing adjusts, or that aid bypasses the budget. Reliable data on net borrowing and lending are not available.[35] The issue of off-budget aid is investigated in Section 2.6. The data on general government revenues as a share of GDP are available. However, in addition to tax revenues and other domestic revenues, these data also include grants, so at least part of on-budget aid is recorded as government revenues. Therefore, as pointed out before, government revenues are not equivalent to the domestic revenues variable from equation (2.1). The results of the impact of aid on government revenues are presented in columns [5] and [6] of Table 2. For both models, with and without lagged government revenues, the estimate of aid's impact is positive. When the lagged dependent variable is included, the impact is statistically different from zero, and 1 pp of aid in GDP increases government revenues by around 0.079 pp. The positive impact of aid on government revenues is expected because part of aid is included in government revenues in the WEO dataset. However, a low coefficient of the impact of aid on government revenues suggests that aid is substituting tax revenues or that it is bypassing the budget.

Since fungibility at the aggregate level has been tested before, the results presented above can be compared to other studies presented in Section 2.2. All studies listed below estimate equation (2.2) (with different controls), so results should be compared to column [2] of Table 2. Since the lagged dependent variable is not included and the error structure is not modeled, it is very likely that these studies suffer from autocorrelation problems. Devarajan et al. (2007) find, for a small sample of 18 sub-Saharan countries in the period 1971–1995, that around 10% of aid is used as tax relief (or bypasses the budget). Basic results of Feyzioglu et al. (1998) confirm that result: in a sample of 14 countries

[35] As discussed in Appendix A.

between 1970 and 1990 only 5% of aid is fungible. However, when the sample size is extended to 38 countries, 67% of aid is fungible, which is closer to the results presented in this study. Chatterjee et al. (2012) is closest to the current study. They find that the coefficient of aid's impact on government expenditures is 0.3,[36] which is slightly higher than the results found in my study for the same equation.

To sum up, fungibility at the aggregate level has been re-examined on a rich panel of 118 countries for the period 1980–2012. Short- and long-term impact of aid on the government expenditures is estimated thanks to the inclusion of the lagged dependent variable and thanks to the use of averaged data. In the long run, aid is partially fungible and increases government expenditures by around 40–50%. The adjustment process is gradual as aid is highly fungible in the short run. Given that the impact of aid on government revenue is lower than that on expenditures, it means that one of the following happens: (i) aid is substituting tax revenues, (ii) the government immediately decreases its own borrowing needs in response to additional aid, or (iii) that aid is bypassing the budget.

2.6 The impact of off- and on-budget aid on government expenditures

Morrissey (2014) starts his review of fungibility and fiscal response studies by observing that the amount of aid reported by donors is substantially higher than the amount spent in recipient countries. This discrepancy can be explained by the large amount of money spent in the donor country on technical assistance. This category of aid includes development consultancy and training, as well as scholarships given to students from developing countries to finance their education in donor countries. There is also a discrepancy between the amount of aid that reaches the recipient country and the amount recorded in the budget. A part of aid directly reaches ultimate beneficiaries in developing countries through e.g. donor operated projects, and thus bypasses the budget.

[36] They find similar results when the difference GMM estimator is used and when an instrument based on geographical and cultural proximity is used. The robustness of these results is tested in chapter three.

Even if aid does not reach the recipient country government's budget, it may still influence government behavior. For example, off-budget aid[37] may finance the projects that would be otherwise financed by government expenditures. The issue of off-budget aid in the context of fungibility has been noted by Devarajan et al. (2007), Holmqvist (2000), McGillivray and Morrissey (2000), and analyzed extensively by Van de Sijpe (2013b).

The presence of off-budget aid should be acknowledged when the impact of total aid on government expenditures is analyzed. The left-hand side of fungibility equations (2.2) and (2.3)—the share of government expenditures (or government revenues) in GDP—includes only on-budget aid, whereas total aid that is on the right-hand side includes both on- and off-budget aid reported by donors. Therefore, if off-budget aid is not accounted for, aid's impact coefficient is expected to be biased downwards since off-budget aid is not directly increasing budget expenditures.[38] Therefore, a marginal effect smaller than one may be a result of the fact that part of aid is not recorded as government spending. Off-budget aid is fungible if it results in a decrease of the government's own spending (when the coefficient next to off-budget aid is smaller than zero). The interpretation of on-budget aid stays the same as for the models with total aid—aid is fungible if one euro of aid results in less than one euro increase in government spending. However, since the presence of off-budget aid is accounted for, fungibility of on-budget aid means that the government has either adjusted domestic revenues or net borrowing.[39]

[37]In this study, the term "off-budget aid" is used to describe any aid disbursement that is not recorded in the recipient country government's budget.

[38]Assuming that aid is not fungible, so that it does not change the way recipient governments spend their own resources, then if all aid is channeled through the budget, one euro of aid increases government expenditures by exactly one euro. However, when half of aid bypasses the budget, then one euro of aid increases government spending only by 50% (so by the amount of on-budget aid), while the other 50% are not recorded in the budget. If no distinction between off- and on-budget aid is made, the estimate of the fungibility coefficient (0.5) will be biased downwards compared to its true value of one. Moreover, Van de Sijpe (2013b) shows in a simple analytical framework that when off-budget aid is not accounted for, estimates of fungibility may be biased.

[39]Van de Sijpe (2013b) also discusses previous studies in the context of off-budget aid. Those relying on aid data provided by donors, i.e. among others McGillivray and Ouattara (2005); Osei et al. (2005); Pettersson (2007a,b), overestimate the extent of fungibility. That argument applies also to the analysis in the previous sections. There are studies (e.g. Pack and Pack, 1990, 1993) that used recipient-based aid data where off-budget aid is treated as an omitted variable. Then, since on- and off-budget aid are generally correlated, the estimate of fungibility of on-budget

Van de Sijpe (2013b) uses technical cooperation (TC) as a proxy for off-budget aid and the value of sector programme aid as a proxy for on-budget aid, and investigates fungibility in health and education sectors.[40] He finds that off-budget aid is, at most, substituting a small part of own government expenditures in the health and education sectors. The results for on-budget aid are inconclusive.

Table 4: Mean of on- and off-budget aid as a percentage of the recipient country's GDP, unweighted (Un) and weighted by GDP (W)

	Total aid		Bilateral aid		Multilateral aid	
	Un	W	Un	W	Un	W
On-budget	7.13	0.97	3.80	0.56	3.06	0.36
Off-budget	2.33	0.28	1.89	0.24	0.43	0.04

Notes: Total aid also includes aid from non-DAC countries

Using Van de Sijpe (2013b)'s methodology, I analyze whether off-budget aid is non-fungible at the aggregate level and investigate whether not accounting for off-budget aid biases fungibility estimates. Technical cooperation is used as a proxy for off-budget aid. In the OECD's DAC2a Table the data on sector programme aid are not available[41] therefore the difference *Total aid – technical cooperation* is taken as a proxy for on-budget aid. It is, admittedly, an imperfect proxy since it includes the off-budget elements as well. Hence, if not accounting for off-budget aid indeed lowers the estimate of aid's impact on government expenditures, the current study's results for fungibility of on-budget aid should be treated as an upper bound. Or, in other words, as a lower bound for the impact of aid on government expenditures.[42]

aid is biased unless the marginal effect of off-budget aid on government expenditures is zero (Van de Sijpe, 2013b).

[40]In the aid-growth literature, Annen and Kosempel (2009) discuss the impact of technical cooperation and non-technical cooperation on growth.

[41]OECD's Credit Report System (CRS) includes data spent on 'sector programme'. However, the CRS dataset is incomplete having many missing values, especially in the 1980s and the 1990s.

[42]Foster (2005) analyzes 14 donors and finds that on average out of any dollar of aid 30 cents are not recorded in the balance of payments, and 20 cents are recorded in the balance but not in the government budget. This means that around 50% of aid is off-budget. Gottret and Schieber (2006) point out that off-budget aid is common in the health sector, with more than 50% of total health spending in Uganda and 46% in Tanzania classified as off-budget.

Table 5: The impact of on- and off-budget aid on government expenditures and revenues

	[1] EXP (OLS)	[2] EXP (FE)	[3] EXP (FE)	[4] EXP (FE)	[5] REV (FE)	[6] REV (FE)
Off-budget aid	2.034**	-0.211	-0.109	0.227	-0.720*	-0.347
	(0.95)	(0.17)	(0.09)	(0.25)	(0.40)	(0.28)
On-budget aid	0.182	0.191***	0.104***	0.106**	0.162***	0.133**
	(0.13)	(0.04)	(0.02)	(0.05)	(0.06)	(0.06)
L.Gov. exp. in GDP			0.619***	0.616***		
			(0.03)	(0.03)		
GDP growth	-0.042	-0.042	-0.009	-0.011	0.100*	0.091**
	(0.08)	(0.04)	(0.03)	(0.03)	(0.05)	(0.04)
L.Inflation	0.006	0.001	0.000	0.000	-0.001	-0.001
	(0.01)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
L.Infant mortality	-0.057*	0.013	-0.001	0.001	-0.005	-0.004
	(0.03)	(0.03)	(0.01)	(0.01)	(0.03)	(0.02)
L.Exp+imp in GDP	0.029	0.010	0.006	0.007	0.008	0.001
	(0.02)	(0.02)	(0.01)	(0.01)	(0.02)	(0.01)
L.Agr. VA	-0.267***	-0.202***	-0.037	-0.042	-0.193**	-0.064
	(0.06)	(0.07)	(0.03)	(0.03)	(0.08)	(0.05)
Off-budget aid sq.				-231.578*		
				(137.93)		
On-budget aid sq.				-1.058		
				(4.45)		
L.Gov. rev. in GDP						0.525***
						(0.09)
Constant	37.403***	27.417***	7.910***	7.489**	25.574***	12.232***
	(5.68)	(4.75)	(2.96)	(2.93)	(5.34)	(3.76)
LR off-budget			-0.286	0.59		
			(0.22)	(0.63)		
LR on-budget			0.273***	0.276*		
			(0.06)	(0.14)		
R-sqr	0.33	0.15	0.49	0.49	0.17	0.36
F-statistic	9.08	12.56	64.68	59.55	5.72	29.75
Obs.	2112	2112	2024	2024	2132	2045

Notes: The dependent variable is the share of government expenditures in GDP (%) (denoted as EXP) or the share of general government revenue in GDP (%) (denoted as REV). *Aid* is measured as the share of aid in GDP (%). Year fixed effects are included in all models (also OLS), country fixed effects are included in the fixed effects model (FE), coefficients are not reported. Standard errors are clustered by country and are reported in parentheses. LR presents the long-run impact of aid on government expenditures from equation (2.4). $^{*}p < 0.1, ^{**}p < 0.05, ^{***}p < 0.01$.

Even when only approximated by the technical cooperation, the share of off-budget aid in GDP is substantial (see Table 4). Technical cooperation accounts for a quarter of total aid, while for bilateral aid this share is even higher: 33%. The large share of off-budget aid indicates that aid bypassing the budget may be an important explanation for partial fungibility of aggregate aid.

Ex ante, it is most likely that off-budget aid is non-fungible. However, it is also possible that partial fungibility or, conversely, the crowding-in effects are present. As technical cooperation constitutes a substantial part of total aid, it may be expected that the government will react to the inflow of off-budget aid (at least to the part that reaches the recipient country) by decreasing own expenditures in similar categories. However, aid illusion regarding off-budget aid may lead to non-fungibility of off-budget aid. While governments are internalizing the expected size and arrival of on-budget aid and treat it (to a large extent) as a substitute for government expenditures, information regarding off-budget aid may not be available, hence government spending is not decreasing. It may also happen that while the government is aware of the inflow of off-budget aid, it spends low amounts of its domestic revenues on expenditures that can be classified as technical cooperation, therefore substantial reduction that would offset inflows of aid is not possible (Van de Sijpe, 2013b). There are even some situations that may lead to the crowding-in effect of off-budget aid. Technical cooperation may be pushing for other types of expenditures. For example, if doctors are taught how to do new diagnostic tests they may simultaneously increase pressure on the government to provide required equipment. Eventually, the fungibility of off-budget aid is an empirical question.

The impact of off-budget and on-budget aid on government expenditures and revenues is presented in Table 5. This table shows that the estimated coefficient of on-budget aid is higher than for total aid, and around 19–27% of on-budget aid is increasing government expenditures. The estimates of the long-run impact are volatile. Each euro of on-budget aid increases government expenditures by 27 cents (for annual data, Table 5) to 59 cents (for the 5 year average, Table 6). As expected, on-budget aid increases government revenue.

The coefficient of off-budget aid is statistically insignificant, but negative in almost all specifications. The interpretation of this insignificant result is

Table 6: The impact of off- and on-budget aid on government expenditures - averaged data

	[1] 3y	[2] 3y	[3] 4y	[4] 4y	[5] 5y	[6] 5y
Off-budget aid	-0.367	-0.013	-0.167	0.158	-0.362	-0.137
	(0.25)	(0.21)	(0.23)	(0.24)	(0.29)	(0.35)
On-budget aid	0.339***	0.239***	0.366***	0.284***	0.490***	0.485***
	(0.07)	(0.04)	(0.09)	(0.07)	(0.11)	(0.12)
L.Gov. exp. in GDP		0.372***		0.307***		0.182**
		(0.06)		(0.08)		(0.08)
GDP growth	-0.114	-0.030	-0.080	-0.040	-0.105	-0.062
	(0.07)	(0.09)	(0.08)	(0.08)	(0.08)	(0.11)
L.Inflation	0.000	0.002	-0.000	-0.001***	-0.001**	-0.001
	(0.00)	(0.01)	(0.00)	(0.00)	(0.00)	(0.00)
L.Infant mortality	0.006	-0.001	-0.005	-0.013	-0.031	-0.035
	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.04)
L.Exp+imp in GDP	-0.004	-0.001	-0.001	-0.001	-0.015	-0.003
	(0.02)	(0.02)	(0.02)	(0.03)	(0.02)	(0.03)
L.Agr. VA	-0.205**	-0.132**	-0.210***	-0.175**	-0.214***	-0.161*
	(0.08)	(0.07)	(0.07)	(0.09)	(0.07)	(0.08)
Constant	32.918***	20.687***	32.774***	21.901***	33.736***	24.309***
	(3.31)	(3.46)	(3.39)	(4.91)	(4.06)	(4.57)
LR off-budget		-0.021		0.22		-0.168
		(0.34)		(0.31)		(0.42)
LR on-budget		0.380***		0.410***		0.592***
		(0.07)		(0.11)		(0.13)
R-sqr	0.20	0.33	0.22	0.32	0.27	0.33
F-statistic	10.99	20.76	11.79	29.66	16.30	14.09
Obs.	734	650	550	474	434	364

Notes: The dependent variable is the share of government expenditures in GDP (%). *Aid* is measured as the share of aid in GDP (%). Year and country fixed effects included in all models, coefficients are not reported. Standard errors are clustered by country and are reported in parentheses. LR presents the long-run impact of aid on government expenditures from equation (2.4). 3y, 4y, 5y means data averaged over 3, 4 and 5 years respectively. $^{*} p < 0.1, ^{**} p < 0.05, ^{***} p < 0.01$.

different than for on-budget aid: the hypothesis of off-budget aid being non-fungible cannot be rejected and the recipient government does not increase government expenditures in response to the inflows of off-budget aid. The negative sign indicates that off-budget aid may be affecting government expenditures and may be (indirectly) substituting some of them. Moreover, the coefficient of the impact of off-budget aid on government revenue is negative (column [5] and [6] of Table 5), which suggests that it may be replacing domestic revenue.

In conclusion, aid recorded in the budget turns out to be partly fungible in the long run. Projects that bypass the recipient country's budget are non-fungible, and there are some signs of a small crowding-out effect. These results suggest that, as expected, when the distinction for on- and off-budget aid is not made, the impact of aid is downward biased (so the fungibility is overestimated). However, the bias is small and statistically insignificant.

2.7 The impact of multilateral and bilateral aid on government expenditures

The extent of fungibility may differ for different types of aid. Many studies focus on sectoral disaggregation of aid and find significant differences in the extent of fungibility. However, to my knowledge recent studies have not included an analysis of disaggregation into bilateral and multilateral aid.

Ram(2003) lists three main differences between bilateral and multilateral aid:

1. Donor motives: Economic and strategic interests are likely to be more important for bilateral aid. For example, starting with the Marshall Plan, US aid closely followed changes in strategic points on the world map, switching its focus to East Asia, Middle East, or Latin America conditional on the geopolitical situation and current foreign policy aims, like the Cold War or the War on Terror (Lebovic, 1988; Todaro and Smith, 2003; Fleck and Kilby, 2010). Japan concentrates its aid on East Asia, particularly neighboring economies that are also major recipients of Japanese foreign direct investments or major trade partners (Berthélemy, 2006; Todaro and Smith, 2003; Ram, 2003). Great Britain and France allocate their aid in their former colonies, while the OPEC supports Arab League countries (Boone, 1996). Berthélemy (2006) finds that all donors (except for Switzerland) allocate more aid to their main trading partners. Younas (2008) confirms this by stating that OECD countries allocate more aid to recipient nations that import goods in which the donor nation has a comparative advantage in production. Alesina and Dollar (2000) summarize evidence for bilateral aid as follows: "*there is considerable* *evidence that the direction of foreign aid is dictated as much by political and strategic considerations, as by the economic needs and policy performance of the recipients. Colonial past and political alliances are major determinants of foreign aid*" (p. 33). Conversely, multilateral aid seems to be more policy and poverty oriented (Maizels and Nissanke, 1984; Dollar and Levin, 2006), and is allocated to countries with good policies (Burnside and Dollar, 2000).

2. Aid conditionalities: Ram (2003) and Berthélemy (2006) point out that multilateral and bilateral aid packages differ in the conditions attached. Multilateral institutions, like the World Bank and the IMF, have conditioned for a long time their aid on so called *structural adjustment and reform programs.* Bilateral donors usually did not use this type of requirement. However, DAC has been working on the harmonization of donor practices for more than 20 years (OECD, 2003), and some bilateral donors follow allocation practices of multilateral organizations, *implicitly* conditioning their aid allocation.

3. Closeness of the relationship between the donor and the recipient: Bilateral donors often have long-lasting relationships with recipients (dating back to colonial times) and therefore often have similar institutions, the same language, a history of personal and commercial interactions, and country-specific knowledge. These factors may facilitate interactions and lead to a better understanding of the recipient country's needs (Ram, 2003 after Cassen and associates, 1994). However, as has been previously argued (Alesina and Dollar, 2000; Burnside and Dollar, 2000), multilateral institutions seem to pay more attention to the recipient country's needs than bilateral donors.

These three main differences suggest important differences between bilateral and multilateral aid, hence even at this level of aggregation the degree of fungibility may be different.[43] *A priori*, it is not clear whether multilateral aid is more or less fungible than bilateral. One of the main reasons of fungibility is that preferences between donors and recipients are not perfectly aligned.

[43] Furthermore, Ram (2003, 2004) analyzed the impact of bilateral and multilateral aid on growth. He showed that both parameters are significant and sizeable, but have opposite signs. This may suggest that the small and sometimes statistically insignificant impact of aid on growth can be decomposed into two statistically significant effects: a strong positive effect of bilateral aid, offset by a strong negative effect of multilateral aid on growth. However, in both studies aid is not instrumented for to account for endogeneity in regressions. Also for the OLS regressions, Burnside and Dollar (2000) find that bilateral aid increases government consumption, while the coefficient of multilateral aid is statistically insignificant.

Table 7: The impact of bilateral and multilateral aid on government expenditures and revenues

	[1] EXP (OLS)	[2] EXP (FE)	[3] EXP (FE)	[4] EXP (FE)	[5] REV (FE)	[6] REV (FE)
Bilateral aid	0.621	0.032	0.034	0.017	-0.058	0.014
	(0.41)	(0.07)	(0.04)	(0.07)	(0.11)	(0.06)
Multilateral aid	0.297	0.321***	0.146***	0.178**	0.259**	0.188
	(0.44)	(0.11)	(0.05)	(0.08)	(0.13)	(0.12)
L.Gov. exp. in GDP			0.621***	0.621***		
			(0.03)	(0.03)		
GDP growth	-0.096	-0.044	-0.009	-0.009	0.106**	0.095**
	(0.09)	(0.04)	(0.03)	(0.03)	(0.05)	(0.04)
L.Inflation	0.003	0.001	0.000	0.000	-0.000	-0.001
	(0.01)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
L.Infant mortality	-0.059*	0.013	-0.001	-0.001	-0.005	-0.004
	(0.03)	(0.03)	(0.01)	(0.01)	(0.03)	(0.02)
L.Exp+imp in GDP	0.038*	0.009	0.005	0.005	0.003	-0.002
	(0.02)	(0.02)	(0.01)	(0.01)	(0.02)	(0.01)
L.Agr. VA	-0.263***	-0.209***	-0.040	-0.040	-0.208***	-0.070
	(0.06)	(0.07)	(0.03)	(0.03)	(0.08)	(0.05)
Bil. aid sq.				3.457		
				(14.09)		
Mul. aid sq.				-9.770		
				(19.89)		
L.Gov. rev. in GDP						0.535***
						(0.09)
Constant	39.278***	27.210***	7.772***	7.797***	25.125***	11.911***
	(6.47)	(4.71)	(2.88)	(2.90)	(5.27)	(3.59)
LR bilateral			0.089	0.05		
			(0.11)	(0.19)		
LR multilateral			0.386***	0.469**		
			(0.13)	(0.22)		
R-sqr	0.27	0.14	0.48	0.48	0.15	0.35
F-statistic	9.99	9.04	62.26	72.52	4.85	36.35
Obs.	2112	2112	2024	2024	2132	2045
Equal coeff.	0.69	0.06	0.18	0.19	0.13	0.30

Notes: The dependent variable is the share of government expenditures in GDP (%) (denoted as EXP) or the share of general government revenue in GDP (%) (denoted as REV). *Aid* is measured as the share of aid in GDP (%). Year fixed effects are included in all models (also OLS), country fixed effects are included in the fixed effects model (FE), coefficients are not reported. Standard errors are clustered by country and are reported in parentheses. LR presents the long-run impact of aid on government expenditures from equation (2.4). $^{*} p < 0.1, ^{**} p < 0.05, ^{***} p < 0.01$.

There are reasons to believe that multilateral aid is less fungible. Under the (idealistic) assumption that recipient governments care about pro-poor actions, since—as argued before—multilateral agencies seem to especially pay attention to poverty, one would expect smaller fungibility of funds from multilateral agencies. Additionally, due to attached conditions that may require a recipient government's own contribution, multilateral funds may be less likely to substitute government spending. However, Morrissey (2004) points out that there is consensus in the literature that aid conditionality does not produce desired effects. Despite the attached conditions governments do not undertake the required reforms. Even if governments are planning to undertake the reforms, imposing conditionality may be unnecessary and possibly even damaging.

Furthermore, bilateral aid is often used as a way to promote products from the donor countries. Tied aid packages require recipients to purchase goods or services produced in donor countries. Around half of the aid channeled to the Least Developed Countries is tied (OECD, 2001). Tied aid tends to favor projects that require capital intensive imports or donor-based technical cooperation, instead of smaller and more poverty-focused programmes, while at the same time being 15–30% more costly for the recipient[44] than untied aid (OECD, 2001). At the sectoral level, if products provided by the donor are not prioritized by the recipient, a diversion of funds is very likely. However, it is not clear what effect tied aid has at the aggregate level. If donors provide products that would not be bought otherwise, then, paradoxically, aid should be non-fungible, since an optimizing government would allocate all other resources as if no aid was provided.

Given the considerations presented above, the assumption of the equality of coefficients of bilateral and multilateral aid can be relaxed. I test for equality of coefficients of bilateral and multilateral using an F-test. Under the null hypothesis that both coefficients are equal the constraint is binding. The p-values of this test are shown in rows *Equal coeff.* in Tables 7 and 8.

The estimated impact of bilateral and multilateral aid on government expenditures is volatile (see Tables 7 and 8). In all fixed effects models multilateral

[44]Compared to the situation in which a recipient can purchase similar products or services at world prices. Effectively, tied aid is a way to subsidize the donor's domestic industry.

Table 8: The impact of bilateral and multilateral aid on government expenditures - averaged data

	[1] 3y	[2] 3y	[3] 4y	[4] 4y	[5] 5y	[6] 5y
Bilateral aid	-0.029	0.092	0.049	0.148	0.231	0.313**
	(0.14)	(0.15)	(0.15)	(0.18)	(0.14)	(0.14)
Multilateral aid	0.642***	0.365**	0.610***	0.414*	0.580**	0.533**
	(0.16)	(0.18)	(0.19)	(0.21)	(0.23)	(0.24)
L.Gov. exp. in GDP		0.371***		0.306***		0.185**
		(0.06)		(0.09)		(0.08)
GDP growth	-0.113	-0.028	-0.071	-0.043	-0.056	-0.028
	(0.07)	(0.08)	(0.08)	(0.08)	(0.08)	(0.11)
L.Inflation	0.000	0.002	-0.001	-0.001***	-0.001**	-0.001*
	(0.00)	(0.01)	(0.00)	(0.00)	(0.00)	(0.00)
L.Infant mortality	0.006	-0.001	0.001	-0.010	-0.029	-0.036
	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.04)
L.Exp+imp in GDP	-0.004	-0.000	0.001	0.001	-0.017	-0.002
	(0.02)	(0.02)	(0.02)	(0.03)	(0.02)	(0.03)
L.Agr. VA	-0.220***	-0.136**	-0.221***	-0.179**	-0.234***	-0.171**
	(0.08)	(0.06)	(0.07)	(0.09)	(0.07)	(0.09)
Constant	33.148***	20.755***	32.595***	21.861***	33.237***	27.633***
	(3.26)	(3.50)	(3.35)	(5.00)	(4.04)	(3.38)
LR bilateral		0.147		0.13		0.384**
		(0.24)		(0.26)		(0.17)
LR multilateral		0.581**		0.596**		0.654**
		(0.27)		(0.30)		(0.28)
R-sqr	0.19	0.32	0.21	0.31	0.25	0.31
F-statistic	10.36	18.32	12.05	28.28	13.56	13.04
Obs.	734	650	550	474	434	364
Equal coeff.	0.02	0.39	0.07	0.47	0.24	0.45

Notes: The dependent variable is the share of government expenditures in GDP (%). *Aid* is measured as the share of aid in GDP (%). Year and country fixed effects included in all models, coefficients are not reported. Standard errors are clustered by country and are reported in parentheses. LR presents the long-run impact of aid on government expenditures from equation (2.4). 3y, 4y, 5y means data averaged over 3, 4 and 5 years respectively. $^{*} p < 0.1, ^{**} p < 0.05, ^{***} p < 0.01$.

aid is statistically different from zero, but coefficients range from 0.32 to 0.64 for models without the lagged dependent variable, and between 0.38 to 0.65 for the long-term impact in the models including the lagged dependent variable. The coefficient of bilateral aid is close to zero and statistically insignificant in the models with annual data (see Table 7), and data averaged over three and four years (see columns [1]–[4] in Table 8). It is statistically significant only for the five year average when lagged government expenditures are included (see column [6] in Table 8). As the standard errors are high, the hypothesis of equal coefficients can be rejected at the 10% level in only three models (column [2] in Table 7, and columns [1] and [3] in Table 8). These results indicate that bilateral aid may be more fungible than multilateral aid.

2.8 Robustness

In order to test the robustness of the results, I discuss outliers, weighting, changes in controls variables, and the impact of military expenditures.[45]

To look for outliers, following Van de Sijpe (2013b), the results are tested by re-estimating the main equations eliminating one country at a time. All values for the estimate of off-budget aid coefficient are stable and between −0.14 and −0.06 (again, statistically insignificant), and for on-budget aid between 0.095 and 0.115. Estimates are also stable for total aid.

This exercise of removing one country at a time was also done at the very initial stage of the research. The Solomon Islands was found to be an outlier and was removed from the sample (details in Appendix B). The impact of the Solomon Islands, a remote country with a population of around half a million people, raises the question whether the results are not driven by small countries that are not representative for the whole sample, and hence are less interesting for donors. There are 25 countries, mostly islands, with an average population below 1 million during the period discussed. Removing them from the sample does not change the results. However, it makes the estimates less precise.

The results are also weighted by either population size or total GDP using the Weighted Least Squares method, where weights are given by the absolute or squared residuals. In both cases, the results stay very close to the original results.

I use lagged control variables (except for growth) arguing that current aid could affect spending through these control variables in the same period. For example, aid may reduce infant mortality, and as a result government

[45]All non-reported results are available upon request.

expenditures will decrease. However, it can also be argued that blocking out this channel (by including current values of control variables) could provide a better estimate of what donors are mostly interested in—the direct effect of aid on fiscal policy choices. As a robustness check, instead of lagged values of the control variables, contemporaneous values are used. The estimates stay very close to the original ones both for the yearly and averaged data. Off-budget aid for averaged data becomes statistically significant (and negative) for the 3- and 5-year average. The impact of aid on government revenues becomes insignificant.

Regressions with additional controls (discussed in Appendix C) have been performed and the results are robust to the changes in the control set.

A particular concern for donors related to sectoral fungibility is a potential diversion of development assistance intended for poverty reduction purposes into military spending. Collier and Hoeffler (2007) point out that as recipient's budgets are partially financed by aid – directly or through fungible projects – aid may be inadvertently financing military expenditures. After instrumenting, they find that aid has a positive and significant effect on military expenditures. Some fungibility studies (Chatterjee et al., 2012) explicitly excluded military expenditures from government expenditures. In the main analysis I take the broadest possible definition of government expenditures (that includes military expenditures) to keep the biggest sample. As a robustness check, the share of military expenditures in GDP is subtracted from the share of government expenditures in GDP, and the newly constructed variable is used as a dependent variable. This decreases the sample size by approximately 30%. Military expenditures are taken from two sources: WEO and from the database of the Stockholm International Peace Research Institute (SIPRI). Table 12 from Appendix E lists results for three dependent variables for a comparable sample: the share of total government expenditures, the share of total government expenditures minus the share of military expenditures according to WEO, and the share of total government expenditures minus the share of military expenditures according to SIPRI. The results for the alternative dependent variables are unchanged compared to the results for the total government expenditures.

2.9 Conclusions

This chapter investigates the impact of aid on government expenditures and government revenues in the short and long run and contributes to the fungibility and fiscal response literature.

Firstly, the question of fungibility at the aggregate level is re-examined using a bigger, more recent, and more balanced dataset than data used in previous studies, taking into account short-term adjustments and long-term impact of aid. Aid is partly fungible in the long run with around 40–50% of aid increasing government expenditures, whereas in the short run approximately 90% of aid substitutes for other sources of government revenue or bypasses the budget. Moreover, there is no evidence for asymmetric response of government expenditures to increases and decreases in aid. The fungibility of aid at the aggregate level means that aid is either substituting for domestic revenues (potentially decreasing taxes), reducing government's net borrowing, or bypassing the budget. Foreign aid inflows increase government revenues, however the coefficient of impact is small (and often insignificant) which suggests that foreign aid substitutes tax revenues.

Secondly, this study investigates the impact of on- and off-budget components on government expenditures, following Van de Sijpe (2013b) who accounted for off-budget aid in the analysis of fungibility in education and health sectors. To my knowledge the distinction between on-budget aid and aid that bypasses the budget has not been done in studies that investigate fungibility at the aggregate level. Technical cooperation is used as a proxy for off-budget aid, and the rest of aid is treated as a proxy of on-budget aid. Off-budget aid is not fungible, and in some regressions there are signs of partial fungibility. Unsurprisingly, on-budget aid is found to be partially fungible.

Thirdly, aid is disaggregated into a bilateral and a multilateral component and the fungibility of both parts is examined. The results suggest that bilateral aid is more fungible than multilateral aid.

Appendix A: Data description

Table 9 describes the variables used in the chapter and their sources. The data were downloaded in April 2014 from: the World Development Indicators (WDI) of the World Bank, the World Economic Outlook (WEO) of the International Monetary Fund, Table 2a of the Development Assistance Committee (DAC) of the Organization for Economic Cooperation and Development (OECD), and the U.S. Census Bureau's International Database (IDB).

The advantage of the DAC2a OECD Table over the Creditor Reporting System (CRS), which is usually used to measure sectoral fungibility, is that DAC aid data are supposed to be more complete,[46] while CRS coverage for the 1980s and the 1990s is low. The DAC2a Table contains data on aid disbursement from bilateral and multilateral DAC members. Additionally, the DAC2a Table provides data on bilateral aid flows from non-DAC countries that agreed to voluntarily report their disbursements and commitments to the OECD. I do not analyze bilateral non-DAC aid due to the following reasons: (i) it is not clear when the reporting process started and how complete the data are, (ii) in the literature, only bilateral aid from DAC countries is used, (iii) the share of non-DAC aid in aid-recipient's GDP is very small compared to bilateral and multilateral DAC aid (only 3% of total aid).

Arndt et al. (2010) argue (on the basis of information obtained from the OECD) that the majority of missing values for aid represents in fact unreported null values. Therefore, I replaced each missing aid value by zero. This enlarges the sample size by approximately 40 observations, but it does not have a significant impact on the results.

General government revenues are described in the WEO dataset as the sum of "taxes, social contributions, grants receivable, and other revenue"(WEO dataset). Hence, it also includes at least the grant element of on-budget aid and is different from the domestic revenues variable in equation (2.1). Data on net borrowing and lending are available in the WEO and WDI datasets. In

[46] However, there are still some problems as noted in OECD (2011): *"While DAC statistics include the outflows from all major multilateral organizations, there is still progress to be made. Data coverage could be improved for UN specialized agencies and trust funds, and the accuracy of sectoral information could be enhanced for a number of UN funds and programmes. The Secretariat is collaborating with the United Nations Department of Economic and Social Affairs (UNDESA) in this respect".*

Table 9: Data sources

Variable	Description	Source
GDP growth	Annual percentage growth rate of GDP	WDI
GDP per capita	GDP per capita, thousands constant 2005$, PPP	WDI
Lit. rate	Literacy rates, adults	WDI
Agr VA	Agricultural value added (% of GDP)	WDI
Dep. ratio	Population aged 65 and above (% of total)	WDI
Pop. growth	Population growth (annual %)	WDI
Exp+imp in GDP	Sum of exports and imports of goods and services as a share of GDP	WDI
Infant mortality	Infant mortality rate, both sexes	IDB
Inflation	Annual inflation rate	WEO
Population	Total population	WEO
Gov. exp. in GDP	Government expenditures as a share of GDP	WEO
Gov. rev. in GDP	Government revenue as a share of GDP	WEO
GDP curr $	GDP in current prices, dollars	WEO
Aid	Net ODA disbursement as a share of GDP	OECD DAC2a
Bil aid	Net ODA disbursement from bilateral donors as a share of GDP	OECD DAC2a
Mul aid	Net ODA disbursement from multilateral donors as a share of GDP	OECD DAC2a
Off-budget aid	Net ODA disbursement for technical cooperation as a share of GDP	OECD DAC2a
On-budget aid	Difference between total aid and technical cooperation as a share of GDP	OECD DAC2a

the WEO dataset, net borrowing and lending is calculated as revenue minus total expenditures. In the WDI dataset, the cash surplus/deficit variable (similar in the description to net borrowing and lending) is also calculated as: “revenue (including grants) minus expense, minus net acquisition of nonfinancial assets”(WDI dataset). Moreover, there are only around 1000 yearly observations. Since for both datasets the real data on borrowing and lending are not available, I do not include them in the analysis. The data on tax revenues as a share of GDP, which could be used instead of government revenues, are available in the WDI dataset. However, the data are incomplete and more than 1000 observations would have to be excluded from the final sample.

Finally, although the infant mortality rate is also available in the WDI database, the data are incomplete. This prompted me to use the U.S. Census Bureau's International Database (IDB) instead.

Appendix B: List of LDCs, LMICs and UMICs

Table 10 presents the list of LDCs, LMICs, and UMICs in 2009 taken from the WDI of the World Bank.[47] The following countries were removed from the sample:

- American Samoa, Bulgaria, Lithuania, Romania, and the Russian Federation, because these countries were not ODA recipients or aid inflows were not reported.
- Cuba, North Korea, Somalia, and the West Bank, because current GDP in the IMF dataset is missing.
- Argentina, Belarus, Brazil, Chile, Iran, Libya, Mexico, and Venezuela, because of the very small aid flows (on average smaller than 0.12% of GDP).
- Guinea-Bissau, Haiti, Iraq, Kosovo, the Marshall Islands, Micronesia, and Tuvalu, due to the missing control variables or missing government expenditures.
- Solomon Island because this country is an outlier.[48]

The final sample consists of 118 countries.

[47] The classification started in 1987. However, in the end of 1980s and at the beginning of 1990s many countries were not classified due to the missing data. The countries that reached the high income status from 1987 include: (i) Central and Eastern European countries that democratized around the year 1990, latter joined the European Union, and some of them became DAC donors; (ii) other EU member states (Cyprus, Greece, Malta, Portugal) and fast growing economies (South Korea, Macao); (iii) relatively small islands (Barbados, New Caledonia, Puerto Rico, Trinidad and Tobago); (iv) resource rich small countries (Equatorial Guinea and Oman). It is unlikely that the exclusion of these countries affects the results, as most of them received on average low aid inflows.

[48] The Solomon Islands has been found to change the estimates, especially for off-budget aid. When the Solomon Islands is in the sample the coefficient of off-budget aid impact is between 0.1 and 0.2 in the fixed effects model. When it is excluded for all the remaining 118 countries it is smaller than -0.1 (in both cases, it is not statistically different from zero).

Table 10: The list of LDCs, LMICs and UMICs

Least Developed Countries and Other Low Income Countries (per capita GNI <= $995 in 2009)	Afghanistan, Bangladesh, Benin, Burkina Faso, Burundi, Cambodia, the Central African Republic, Chad, Comoros, the Democratic Republic of the Congo, Eritrea, Ethiopia, the Gambia, Ghana, Guinea, Guinea-Bissau, Haiti, Kenya, the Democratic Republic of Korea, the Kyrgyz Republic, the Lao PDR, Liberia, Madagascar, Malawi, Mali, Mauritania, Mozambique, Myanmar, Nepal, Niger, Rwanda, Sierra Leone, the Solomon Islands, Somalia, Tajikistan, Tanzania, Togo, Uganda, Zambia, Zimbabwe.
Lower Middle Income Countries and Territories (per capita GNI $996–$3 945 in 2009)	Angola, Armenia, Belize, Bhutan, Bolivia, Cabo Verde, Cameroon, China, the Republic of the Congo, Cote d'Ivoire, Djibouti, Ecuador, the Arab Republic of Egypt, El Salvador, Georgia, Guatemala, Guyana, Honduras, India, Indonesia, Iraq, Jordan, Kiribati, Kosovo, Lesotho, Maldives, the Marshall Islands, the Federated States of Micronesia, Moldova, Mongolia, Morocco, Nicaragua, Nigeria, Pakistan, Papua New Guinea, Paraguay, the Philippines, Samoa, Sao Tome and Principe, Senegal, Sri Lanka, Sudan, Swaziland, the Syrian Arab Republic, Thailand, Timor-Leste, Tonga, Tunisia, Turkmenistan, Tuvalu, Ukraine, Uzbekistan, Vanuatu, Vietnam, the West Bank and Gaza, the Republic of Yemen.
Upper and Middle Income Countries (per capita GNI $3 946–$11 905 in 2009)	Albania, Algeria, American Samoa, Antigua and Barbuda, Argentina, Azerbaijan, Belarus, Bosnia and Herzegovina, Botswana, Brazil, Bulgaria, Chile, Colombia, Costa Rica, Cuba, Dominica, the Dominican Republic, Fiji, Gabon, Grenada, the Islamic Republic of Iran, Jamaica, Kazakhstan, Lebanon, Libya, Lithuania, Macedonia FYR, Malaysia, Mauritius, Mexico, Montenegro, Namibia, Palau, Panama, Peru, Romania, the Russian Federation, Serbia, Seychelles, South Africa, St. Kitts and Nevis, St. Lucia, Vincent and the Grenadines, Suriname, Turkey, Uruguay, the Bolivarian Republic of Venezuela.

Appendix C: Discussion of control variables

In addition to the agricultural value added that is used to control for spending in the agriculture sector and for the level of development, Chatterjee et al. (2012) suggest that the real per capita GDP proxies for income, and thus can be used to control for the size of the government. Feyzioglu et al. (1998) refer to Wagner's law which states that development is accompanied by an increase of the share of government expenditures in GDP. Moreover, according to Van de Sijpe (2013b), aid expressed as a share of GDP is very likely to be correlated with GDP per capita. Additionally, Feyzioglu et al. (1998) point out that per capita GDP is correlated with the agriculture share in GDP, infant mortality rate, and school enrollment (so also with literacy rate). As a result the estimated coefficients of these variables may be affected if GDP per capita is not controlled for. However, this variable is likely to be non-stationary. The results do not change when per capita GDP or growth (which is first-differenced GDP) is included in the regressions.

Population growth and the dependency ratio are tested as proxies for the health-care and social security spending (in addition to the infant mortality rate), while the literacy rate is used as a control for expenditures in the education sector. Population growth and the dependency ratio were found to be statistically insignificant at a 5% level and they neither affect the estimates of interest nor the precision of estimates. Therefore, both are excluded from the reminder of the analysis. Additionally, the Variance Inflation Factor (VIF) for projection of dependency ratio on the other variables equals 49, which is above the usually assumed cut-off values proposed in the literature (5 or 10). Therefore, the dependency ratio is not included as a regressor. The literacy rate among adults affects results. However, this is likely caused by the fact that for many countries this rate is almost constant over time. When the literacy rate is not controlled for, it is captured by the fixed effects, which causes it to be highly collinear with the recipient's fixed effects (VIF equals 270). Therefore, the literacy rate is not included as a regressor.

Appendix D: Asymmetric response to decreases and increases

Table 11: The impact of aid on government expenditures for increases and decreases of aid

	(1) Increase	(2) Increase	(3) Decrease	(4) Decrease	(5) Both	(6) Both
Total aid	0.209***	0.091**	0.146***	0.079***		
	(0.08)	(0.04)	(0.05)	(0.02)		
L.Gov. exp. in GDP		0.620***		0.626***		0.627***
		(0.03)		(0.03)		(0.03)
Aid change(decrease)			-0.046	0.057		
			(0.07)	(0.06)		
Aid change(increase)	-0.166*	-0.024				
	(0.09)	(0.06)				
Aid when decreasing					0.167**	0.050
					(0.07)	(0.03)
Aid when increasing					0.143***	0.083***
					(0.05)	(0.02)
GDP growth	-0.047	-0.008	-0.037	-0.003	-0.038	-0.001
	(0.04)	(0.03)	(0.04)	(0.03)	(0.04)	(0.03)
L.Inflation	0.002	0.000	0.002	0.000	0.002	0.000
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
L.Infant mortality	0.016	-0.000	0.013	-0.001	0.013	-0.001
	(0.03)	(0.01)	(0.03)	(0.01)	(0.03)	(0.01)
L.Exp+imp in GDP	0.005	0.004	0.007	0.005	0.007	0.005
	(0.02)	(0.01)	(0.02)	(0.01)	(0.02)	(0.01)
L.Agr. VA	-0.205***	-0.040	-0.208***	-0.040	-0.208***	-0.040
	(0.07)	(0.03)	(0.07)	(0.03)	(0.07)	(0.03)
Constant	26.682***	7.685***	27.138***	7.610***	27.110***	7.631***
	(4.63)	(2.86)	(4.68)	(2.83)	(4.66)	(2.82)
R-sqr	0.15	0.48	0.14	0.48	0.14	0.49
F-statistic	7.66	62.91	8.16	60.00	8.70	63.36
Obs.	2112	2024	2112	2024	2112	2024

Notes: The dependent variable is the share of government expenditures in GDP. *Total aid* is measured as the share of aid in GDP (%). Aid change (increase) is equal to $I_t(AID_{it} - AID_{it-1})$ from equation (2.5), and aid change (decrease) to $D_t(AID_{it} - AID_{it-1})$ from equation (2.6). Aid when decreasing is equal to $I_t AID_{it}$, and Aid when decreasing to $D_t AID_{it}$, both from equation (2.7). Year and country fixed effects are included in all models, coefficients are not reported. Standard errors are clustered by country, reported in parentheses. $^{*} p < 0.1,^{**} p < 0.05,^{***} p < 0.01$.

Appendix E: Military expenditures

Table 12: The impact of total aid on government expenditures minus military expenditures

	(1) EXP	(2) EXP	(3) WO	(4) WEO	(5) SIPRI	(6) SIPRI
Total aid	0.157***	0.094***	0.159***	0.098***	0.159***	0.102***
	(0.03)	(0.02)	(0.04)	(0.02)	(0.04)	(0.02)
L.Gov. exp. in GDP		0.628***		0.568***		0.560***
		(0.05)		(0.05)		(0.05)
GDP growth	-0.023	0.009	-0.003	0.030	-0.005	0.025
	(0.05)	(0.04)	(0.04)	(0.03)	(0.04)	(0.03)
L.Inflation	-0.003	0.000	-0.001	0.000	-0.002	0.000
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
L.Infant mortality	-0.017	-0.013	-0.023	-0.021	-0.032	-0.028
	(0.04)	(0.02)	(0.03)	(0.02)	(0.03)	(0.02)
L.Exp+imp in GDP	0.017	0.010	0.014	0.011	0.017	0.012
	(0.02)	(0.01)	(0.02)	(0.01)	(0.02)	(0.01)
L.Agr. VA	-0.172**	-0.026	-0.197***	-0.049	-0.183***	-0.053*
	(0.07)	(0.03)	(0.06)	(0.03)	(0.06)	(0.03)
Constant	27.704***	9.689***	25.924***	8.731***	25.293***	9.229***
	(5.37)	(2.26)	(4.47)	(2.18)	(4.77)	(2.16)
R-sqr	0.20	0.52	0.23	0.54	0.25	0.53
F-statistic	6.78	42.21	6.04	30.01	6.61	28.86
Obs.	1465	1465	1527	1465	1463	1463

Notes: The dependent variable is the share of government expenditures in GDP (%): denoted as SIPRI (when the SIPRI database is used) and WEO for the WEO database. *Total aid* is measured as the share of aid in GDP (%). Year and country fixed effects are included in all models, coefficients are not reported. Standard errors are clustered by country and are reported in parentheses. $^{*} p < 0.1, ^{**} p < 0.05, ^{***} p < 0.01$.

Table 13: The impact of off- and on-budget aid on government expenditures minus military expenditures

	(1) EXP	(2) EXP	(3) WO	(4) WEO	(5) SIPRI	(6) SIPRI
Off-budget aid	-0.342	-0.285**	-0.252	-0.210	-0.191	-0.145
	(0.24)	(0.13)	(0.24)	(0.15)	(0.22)	(0.14)
On-budget aid	0.192***	0.121***	0.188***	0.120***	0.183***	0.119***
	(0.04)	(0.02)	(0.04)	(0.02)	(0.04)	(0.02)
L.Gov. exp. in GDP		0.625***		0.566***		0.558***
		(0.05)		(0.05)		(0.05)
GDP growth	-0.031	0.002	-0.010	0.024	-0.011	0.021
	(0.05)	(0.04)	(0.04)	(0.03)	(0.04)	(0.03)
L.Inflation	-0.003	-0.000	-0.002	0.000	-0.002	0.000
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
L.Infant mortality	-0.017	-0.013	-0.022	-0.021	-0.032	-0.028
	(0.04)	(0.02)	(0.03)	(0.02)	(0.03)	(0.02)
L.Exp+imp in GDP	0.018	0.012	0.016	0.012	0.018	0.013
	(0.02)	(0.01)	(0.02)	(0.01)	(0.02)	(0.01)
L.Agr. VA	-0.159**	-0.017	-0.186***	-0.042	-0.174***	-0.047*
	(0.07)	(0.03)	(0.06)	(0.03)	(0.06)	(0.03)
Constant	28.280***	10.223***	26.351***	9.165***	25.711***	9.584***
	(5.37)	(2.18)	(4.46)	(2.05)	(4.76)	(2.07)
R-sqr	0.21	0.53	0.24	0.54	0.25	0.54
F-statistic	11.89	41.00	10.74	27.40	12.73	27.86
Obs.	1465	1465	1527	1465	1463	1463

Notes: The dependent variable is the share of government expenditures in GDP (%): denoted as SIPRI (when the SIPRI database is used) and WEO for the WEO database. *Total aid* is measured as the share of aid in GDP (%). Year and country fixed effects are included in all models, coefficients are not reported. Standard errors are clustered by country and are reported in parentheses. $^{*} p < 0.1, ^{**} p < 0.05, ^{***} p < 0.01$.

CHAPTER 3

Instrumental Variables and the Impact of Aid on Government Expenditures

One of the main problems of estimating the causal effects of aid on development indicators is endogeneity of aid in these regressions. This article implements and tests the robustness of two instrumental variables methods used in the studies of fungibility at the aggregate level. First, the difference and system GMM estimators are discussed and implemented on two distinct datasets. The results are imprecise and not robust to the changes in specification. Second, an instrument set based on cultural and historical proximity between the donor and the recipient country is used to account for potential endogeneity of aid. The specification tests show that this instrument set is weak. Overall, neither the GMM estimators nor the proposed IV estimator are able to establish a causal relationship between aid and government expenditures.

I am grateful to Chris Elbers and Jan Willem Gunning for abundant comments and suggestions, to Artūras Juodis for advice on the dynamic panel estimators, to Jukka Pirttilä for useful suggestions during my stay at UNU-WIDER that inspired this chapter, and to Mark Kagan and Paul Muller for suggestions and other help. I am grateful to Santanu Chatterjee, Paola Giulianoy, and Ilker Kaya, as well as to Karen L. Remmer, for sharing their codes and datasets. The usual disclaimers apply.

3.1 Introduction

The effectiveness of foreign aid[49] remains a topic of great public interest. There is also an ongoing debate about the methods that are used to asses the impact of aid on various development indicators: growth, poverty reduction or education. One of the main problems that researchers struggle with is endogeneity of aid in regressions explaining these indicators. Various instruments have been proposed to account for endogeneity. Both difference and system GMM estimators, that use internal instruments, have been applied by Alvi and Senbeta (2012); Annen and Kosempel (2009); Rajan and Subramanian (2008), among others. Boone (1996) proposed a popular set of instruments based on recipients' characteristics. Various modifications of this instrument set, for example including donor-recipient interactions, have been employed by Burnside and Dollar (2000); Hansen and Tarp (2000, 2001); Dalgaard et al. (2004); Rajan and Subramanian (2008); Arndt et al. (2010, 2014), and many others. Also, quasi-experiments have been used to determine the impact of aid on growth (Galiani et al., 2014).

In spite of enormous efforts to find a strong and valid instrument, it seems that the quest is not finished. For example, Deaton (2010) points out that the size of population, an instrument used in many studies, is not exogenous to growth, whereas, as Clemens et al. (2012) note, this instrument accounts for almost all the instruments' strength in Boone (1996) and Rajan and Subramanian (2008).

Endogeneity of aid may also be a problem for estimating the causal impact of aid on government expenditures. The recipient country's government may be able to determine the timing of aid's disbursement, and then aid may be determined simultaneously with government expenditures. Aid may also be channeled to countries that fail to provide merit goods. In the limited literature on fungibility at the aggregate level[50] two methods have been used to account for endogeneity. Chatterjee et al. (2012) and Remmer (2004) use the difference and system GMM estimators as a robustness check to

[49] This chapter uses standard terminology in the aid effectiveness literature. Aid is defined as the Official Development Assistance.

[50] Aid is said to be fungible at the aggregate level if it raises government expenditure by less than the full amount.

non-instrumental methods,[51] and Chatterjee et al. (2012) use cultural and geographical characteristics to perform instrumental variable regressions. The aim of the current chapter is to test the robustness of these methods in the context of fungibility at the aggregate level.

The difference and system GMM estimators have been developed by Holtz-Eakin et al. (1988); Arellano and Bond (1991); Arellano and Bover (1995); Blundell and Bond (1998). For practitioners, three aspects of the dynamic panel estimators have been crucial and can explain their popularity: the abundance of potential internal instruments, correcting the dynamic panel bias, and handling fixed effects. What is more, both estimators have been applied in popular econometric software such as *Stata*, which facilitates but also automates their use. In the article that introduces one of the most popular commands to estimate these models, Roodman (2009a) explicitly warns that the “black box” that these commands create increases the risk of unwittingly misusing these complicated estimators. I follow the instructions provided by Roodman (2009a) and employ the difference and system GMM estimators to investigate the impact of aid on government expenditures. I report and discuss all specification choices, and test sensitivity of the results. Two datasets are used. First, I build upon my previous analysis in chapter two, using the same extensive dataset of 118 developing countries for the period 1980–2012. Second, I use the dataset provided by Chatterjee et al. (2012) that includes 67 countries for years 1972–2000 and test the robustness of their findings.

The studies that investigate fungibility at the aggregate level usually use panel data. Panel data allow to decrease the bias coming from heterogeneity and to control for endogeneity related to fixed effects such as partially time-invariant nonrandom assignment of aid. However, they also increase the difficulty of finding instruments to account for endogeneity of foreign aid in regressions, because these instruments have to be time-varying. Chatterjee et al. (2012) use the instrument set proposed by Tavares (2003) and Larraín and Tavares (2004). I construct a similar instrument set and apply it to the dataset from chapter two, as this dataset differs considerably from the dataset used by Chatterjee et al. (2012).

[51] Both estimators are also employed in the preceding working paper (Marć, 2012).

The chapter proceeds as follows. The next section discusses and implements the difference and system GMM estimators. In the third section, an instrument based on geographical and cultural proximity is used to estimate the impact of aid on government expenditures. The last section concludes.

3.2 System and difference GMM estimators

The difference and system GMM estimators have become a popular method to account for endogeneity of regressors and dynamic panel bias caused by the inclusion of a lagged dependent variable. Chatterjee et al. (2012) use the difference GMM as a robustness check to investigate fungibility at the aggregate level and report similar results to their fixed effects regressions, IV regressions, and to the unreported system GMM estimation.[52] Remmer (2004) mentions GMM methods in her robustness section. In this section, I discuss and apply both difference and system estimators to investigate whether they can be helpful in estimating the causal effect of development assistance on government expenditures. Two distinct data sources are used: the dataset from chapter two and the dataset received from Chatterjee et al. (2012).

Methodological introduction

Dynamic panel estimators have become increasingly popular because they allow the dependent variable to be dynamic and dependent on past realizations, independent variables that are correlated with past and current realizations of the error (so variables that are not strictly exogenous), fixed individual effects, and heteroscedasticity and autocorrelation over time but not across units (Roodman, 2009a).

The difference GMM, developed by Arellano and Bond (1991) and Holtz-Eakin et al. (1988), takes first-differences to remove unobserved time-invariant unit-specific effects. The lags of levels are used to instrument the right-hand side (differenced) variables under the assumption of no serial correlation. However, for the persistent times series when the number of observations is small, the instrumental variables are often weak and therefore are poor

[52] In the footnote, these authors acknowledge that the methods are not without problems.

instruments for the first differenced variables (Bond et al., 2001). To account for this problem, Arellano and Bover (1995) and Blundell and Bond (1998) proposed a system estimator. The system GMM improves upon the Arellano-Bond estimator by additionally including moment conditions in terms of levels: lagged differences are used to instrument levels. This allows more instruments to be introduced, and hence improves efficiency. However, it comes at the cost of an additional assumption: that first differences of instrument variables are uncorrelated with the fixed effects.

The first-difference transformation of data is commonly used in the empirical literature. In addition to that, Arellano and Bover (1995) proposed a "forward orthogonal deviations" (FOD) transformation. Instead of taking the first difference, this transformation subtracts the mean of the remaining future observations available in the sample from the observation. It is especially appropriate for datasets that contain many gaps, as fewer observations are lost compared to the difference transformation. For balanced panels, both transformations return numerically identical estimations.[53]

The GMM estimators are especially designed for datasets that have many panels and few periods: so called "large N, small T" situations. Bun and Kiviet (2006) and Roodman (2009a,b) point out that the number of instruments increases quadratically in the time dimension, which is especially problematic for long panels. Numerous instruments can overfit instrumented regressors and fail to remove the endogenous component. This leads to a bias as the estimates converge towards non-instrumented estimators. In the literature on fungibility at the aggregate level mostly long panels (20–30 years) of annual data are used (e.g. chapter two, Chatterjee et al. 2012; Feyzioglu et al. 1998; Remmer 2004), hence the problem of too many instruments is pronounced.[54] Therefore, when investigating fungibility at the aggregate level, researchers have to be careful when applying the GMM estimators and should be clear on exact specification choices.

[53] Hayakawa (2009) shows in the Monte Carlo study that this estimator tends to work better than the first-difference estimator. However, he notes that: "As T gets larger, the GMM estimator of autoregressive coefficient from FOD model tends to perform better than that from DIF model. However, for other coefficients, this tendency is not always true." (p. 6).

[54] Chapter two uses also data averaged over three, four and five years. Then the number of periods goes down to eleven, eight and six respectively, which greatly reduces the number of instruments.

There is no consensus on what number of instruments causes the problems mentioned. Roodman notes that the threshold of 100 implemented in *Stata* software is generous and he suggests as a preliminary rule of thumb that the number of instruments should not exceed the number of units. This creates a problem for most models in the fungibility literature, as they use annual data and long panels. There are two techniques used to reduce the number of instruments while attempting to maintain necessary information: collapsing the instrument set or limiting the number of lags that can be used as instruments.[55] Both make the instrument count linear in T. Nonetheless, for longer panels the number of instruments can still be above the generous threshold, especially if not only lags of the lagged dependent variable, but also lags of endogenous and predetermined variables are used as instrument. If both techniques are combined, the instrument set is independent of T.

All these methods require the researcher to decide on an appropriate transformation of the instrument set. Mehrhoff (2009) suggests taking it a step further and "letting the data decide". The Principal Component Analysis (PCA) can be applied to the instrument set to condense the information content of the abundant internal instrument set to a lower number of instruments. Using Monte Carlo simulations, Mehrhoff (2009) shows that the GMM estimator with a factorized instrument set that is both collapsed and lag limited has both a lower bias and a lower root mean squared error than any of the alternatives offered above. He also points out that the next best approach (close to the PCA) is to both collapse and lag limit the instrument set. The most frequently used techniques in applied research, the instrument set that is not transformed or the instrument set with the lag limited to one, are the worst choices.[56] However, it is not a method firmly established in the literature, therefore I use it as an additional robustness test.

Conversely, while long panels bring problems for the proper use of the instrument set, they decrease the dynamic panel bias. Nickell (1981) points out that the demanding process of mean subtraction in within estimator creates a correlation between the regressor and the error term, which causes the

[55] For a reference see Roodman (2009a,b); Mehrhoff (2009).

[56] Factor analysis in GMM settings is also discussed by Bai and Ng (2010) and Kapetanios and Marcellino (2010)

dynamic panel bias. This bias is especially severe for small values of T: for the number of panels going to infinity, it equals $(\widehat{\rho}\text{-}\rho)$[57]= $-(1+\rho)/2$ for $T=2$ and $-(1+\rho)(2+\rho)/2$ for $T=3$. The bias decreases for larger values of T and for reasonably large values of T the bias approximates to $(\widehat{\rho}\text{-}\rho)\simeq -(1+\rho)/(T-1)$, so the persistence of y is always underestimated for $\rho > 0$. As an example, for $T=10$ and $\rho=0.5$ the bias is -0.167, and for T around 30 it is three times smaller. Therefore for long panels, if researchers are mostly worried about dynamic panel bias and not about endogeneity of regressors, Roodman (2009a) suggests using the less demanding fixed effects estimators.[58] Otherwise, both difference and system estimators account for the bias.[59]

The Hansen and Sargan tests are canonical checks for instrument validity. For an overidentified equation, they test the null hypothesis of joint validity of all instruments. However, as Roodman (2009a,b) points out the Hansen test is not without drawbacks. He discusses Monte Carlo simulations done by Bowsher (2002) and Andersen and Sørensen (1996). They document that the proliferation of instruments greatly weakens the Hansen test. Roodman (2009a,b) warns that very high values of the Hansen test should be always approached with caution. The difference-in-Sargan and difference-in-Hansen tests are used to check the validity of a subset of instruments. For example, both tests can be used to compare difference and system GMM results, testing the validity of additional instruments introduced by level equations. Both the Sargan and difference-in-Sargan tests are not as vulnerable to a large number of instruments as Hansen's tests, however they require homoscedastic errors. Therefore, a rejection of the null hypothesis for these tests may as well mean that the assumption of the time-series and cross-sectional homoscedasticity is violated.

On the top of the problems listed above, even appropriately applied estimators raise concerns. The system GMM estimator has been introduced because of weak instrument bias in the difference estimator. However, as Bun and Windmeijer (2010) show, the system estimator may also suffer from

[57] $\widehat{\rho}$ is the estimated coefficient of the lagged dependent variable, ρ is the true value.

[58] See the analysis in chapter two in the context of fungibility at the aggregate level.

[59] The literature focuses on the bias caused by the inclusion of a lagged dependent variable. Juodis (2013) shows that the direction of bias of other regressors is ambiguous, and depends on nuisance parameters and the process of data.

weak instrument problems and it should not be taken for granted that the instruments are strong. Also, the stationarity assumption in the system GMM is difficult to test and may be invalid. Arndt et al. (2010) note that for the aid-growth regressions this assumption is unlikely to hold because of the high complexity of the growth process and because “the country fixed effects are expected to incorporate determinants of steady-state income levels that may correlate with growth along individual countries’ steady-state transition paths” (p. 4).[60]

Methodology

In applying the difference and system GMM estimators, this chapter carefully follows the instructions from “How to Do xtabond2: An Introduction to Difference and System GMM in Stata” by Roodman (2009a). Roodman concludes his article with the following recommendation: “Report all specification choices. Using these estimators involves many choices, and researchers should report the ones they make—difference or system GMM; first difference or orthogonal deviations; one- or two-step estimation; nonrobust, cluster robust, or Windmeijer-corrected cluster-robust errors; and the choice of instrumenting variables and lags used” (p. 129). I follow this advice and to increase the clarity, appropriate *Stata* commands are referred to in brackets in italics.

Both GMM estimators are first applied to the dataset from chapter two that investigates the impact of aid on government expenditures. In this study, two econometric models are used. First, a basic fixed effects model, similar to other models used in the fungibility literature (e.g. Chatterjee et al. 2012), is used:

$$GovExp_{it} = \alpha AID_{it} + \beta X_{it} + \gamma_t + \delta_i + \varepsilon_{it} \tag{3.1}$$

where $GovExp_{it}$ represents the share of government expenditures in GDP and AID_{it} the share of the Official Development Assistance in GDP. X_{it} is a set of controls, γ_t are year fixed effects, and δ_i represents the recipient’s country fixed effects. ε_{it} is an error term. Subscript ‘i’ refers to a country and ‘t’ to time. In this model, α is the coefficient of main interest: if α is equal to 1, aid

[60] Bond et al. (2001) and Hauk and Wacziarg (2009) discuss more extensively GMM models in the context of growth theory.

is non-fungible. α less than 1 but greater than 0 means that aid is partially fungible. α equal to 0 indicates that aid is fully fungible.

Second, because of the likely dynamic misspecification, the model that includes the lagged dependent variable is also used:

$$GovExp_{it} = \rho GovExp_{it-1} + \alpha AID_{it} + \beta X_{it} + \gamma_t + \delta_i + \varepsilon_{it} \tag{3.2}$$

The dataset consists of 118 countries in years 1980–2012, which gives around 2050 observations for annual data. Both models are estimated using annual data. Data averaged over three, four and five years are also used to account for the business cycle and measurement error. The averages allow to estimate the long-run impact of aid on government expenditures. The following controls are included: growth, lagged agricultural value added as a share of GDP, lagged sum of imports and exports as a share of GDP, lagged infant mortality rate and lagged inflation. To investigate the impact of aid on government expenditures in the current study, I use equation (3.2), as the difference and system GMM estimators are foremost designed for dynamic panel data. Equation (3.1) is used to replicate the results from Chatterjee et al. (2012) article.

The system estimator is preferred over the difference estimator because the time series are likely to be persistent. The difference estimator is used as a robustness check. For both estimators, the most commonly used difference transformation of the data is used. The forward orthogonal deviations (FOD) transformation (in *Stata*: *orthogonal*) is used only as a robustness check because the dataset does not contain many gaps (the results are available upon request).

Based on findings by Windmeijer (2005) I use the two-step estimator with Windmeijer's correction (*twostep robust*) that was found to lead to a more accurate inference than the the one-step estimation.

To account for correlation across individuals, time dummies are included. All regressors are put into the instrument matrix:

- Time dummies are treated as strictly exogenous, so they are inserted as one column for each time dummy (*ivstyle()*).
- Aid and growth are treated as endogenous, therefore lags two and longer are used as instruments (*gmmstyle()*).

- Lagged government expenditures are predetermined, therefore lags one and longer are used as instruments (*gmmstyle()*).
- Lagged control variables are treated as either strictly exogenous (*ivstyle()*), or predetermined (*gmmstyle()*).

As has been indicated before, long panels are prone to the problem of too many instruments. The instrument set is collapsed (*coll*) and/or lagged limited (*lag*), and the results are tested for sensitivity to both reducing and extending the instrument set. Also, as suggested by Mehrhoff (2009), the Principal Component Analysis (*pca*)[61] is used to extract essential information and choose the optimal number of instruments.

To illustrate the problem of too many instruments, I discuss equation (3.2) with the assumptions about the instrument matrix listed above for the dataset from chapter two. There are seven right-hand side variables: aid and growth (both treated as endogenous), lagged government expenditures (predetermined) and four lagged control variables (for simplicity, I assume that all are predetermined). So there are seven variables to instrument in GMM-fashion. If the instrument set is collapsed, it means that for each time period there is one instrument for each variable, which gives roughly 220 instruments. Additionally, time dummies are treated as strictly exogenous, which increases the instrument count by 33. In total, there are more than 250 instruments. If the instrument set is not collapsed, but lags are limited to one, the problem is even more pronounced and the instrument set equals 450 instruments. As has been indicated by Mehrhoff (2009), collapsing and lag limiting the instrument set at the same time is a superior solution to each of these options. It limits the instrument count to between 40 and 75, depending on the exact number of lags used for the annual data. Thus, it gives hopes that the Hansen test of overidentification restriction is more reliable.

Both system and difference GMM estimators are applied to the dataset used in chapter two. Aid and growth are always treated as endogenous and lagged government expenditures are treated as predetermined. Seven models

[61] In *Stata* it has been recently implemented into Rooodman's *xtabond2* command (package st0159_1). The *pca* option replaces gmm-style instruments with their principal components. By default, only components with eigenvalues of one and above are retained, unless more are needed to prevent underidentification. It is possible to override this setting using the option *components()*.

are estimated for the annual data, and the first six for the averaged data. The additional methodological choices are described below:

1. The system estimator, instrument set collapsed, lagged control variables predetermined.
2. The system estimator, instrument set collapsed and lag limited to one, lagged control variables predetermined.
3. As above, but lag limited to three.
4. As above, but lag limited to five.
5. The system estimator, instrument set collapsed and lag limited to one, lagged controls strictly exogenous.
6. The difference estimator, instrument set collapsed and lag limited to one, lagged control variables predetermined.
7. The system estimator, instrument set collapsed and lag limited to one, lagged control variables predetermined, only data for years 2006–2012 used.

The first model uses a large number of instruments, as the instrument set is only collapsed. The second model is the preferred one, and the following two specifications test the robustness of results to the extension of the instrument set. In model five, I assume that lagged control variables are strictly exogenous. Admittedly, strict exogeneity is unlikely. However, I include this specification to check whether the results are robust to change in the assumptions regarding nuisance variables. In the sixth specification, the difference estimator is used as a robustness check to the preferred specification number two. The last model uses the preferred specification, but looks at only the seven most recent years. There are two reasons to include this specification as a robustness check. Firstly, violation of the stationarity condition is more likely for the initial periods, and this assumption is more likely to be satisfied for the recent periods. Secondly, for shorter panels the risk of structural breaks is smaller. The seventh specification is not estimated for the averaged data, as the number of periods is too small.

Results based on the dataset from chapter two

Table 14 presents the results for the annual data for the seven specification choices that have been listed before. There is no autocorrelation according to the Arellano-Bond autocorrelation test. In all models the estimate of the lagged dependent variable is similar and always statistically significant. The Hansen test (that can be weakened by a large number of instruments) does not reject the null of instrument validity. As Roodman (2009a) warns, it gives a very optimistic value in column [1] that has the highest number of instruments. However, the Sargan test (that is robust to the number of instruments) raises concerns regarding the instrument validity (or to the homoscedasticity assumption) as the null hypothesis is rejected. The coefficient of aid impact is statistically insignificant and for some models it even changes its sign.

In the next step, the system and difference GMM estimators are applied to the data averaged over four years and the results are presented in Table 15. Taking the average of data means that the number of periods decreases to eight, and creates the "small T, big N" situation that these estimators are designed for. It greatly reduces the number of instruments, reducing the risk of overfitting. For some specifications, the Arellano-Bond autocorrelation test fails at rejecting the null hypothesis of autocorrelation. The results of both Hansen and Sargan tests are volatile, for example increasing the number of lags used from three (specification [3]) to five ([4]) changes the conclusion on the instrument validity for the Sargan test. The coefficient of the lagged dependent variable is very high, often close to one. The estimate of the variable of main interest, aid, is insignificant at the 5% level.

The results for the data averaged over three and five years are presented in Appendix A in Tables 19 and 21. For the three year average the coefficient of aid is always insignificant, whereas the coefficient of the lagged dependent expenditures is significant, but the estimate is very volatile. For the five year average, aid appears to have a significant impact on government expenditures in some specifications, but negative and insignificant in others. The estimate of the lagged dependent variable is unstable, and sometimes even higher than one. Relatively small changes in the specification, like increasing the lag limit from one to three, substantially change the results.

Following the advice of Mehrhoff (2009), the Principal Component Analysis is also used to limit the instrument set. Aid and growth are always treated as endogenous and the lagged government expenditures are treated as predetermined. Lagged control variables are predetermined, except for one model. Six models are estimated with the methodological choices described below:

Table 14: Results for the system and difference GMM estimator

	[1]	[2]	[3]	[4]	[5]	[6]	[7]
Total aid	0.031	0.002	-0.022	-0.012	0.166	0.043	0.035
	(0.05)	(0.13)	(0.06)	(0.06)	(0.21)	(0.07)	(0.05)
L.Gov. exp. in GDP	0.511***	0.497***	0.530***	0.568***	0.517***	0.467***	0.507***
	(0.06)	(0.07)	(0.07)	(0.07)	(0.08)	(0.07)	(0.12)
GDP growth	-0.068	-0.087	0.110	0.093	-0.025	0.237*	0.172
	(0.07)	(0.16)	(0.10)	(0.07)	(0.16)	(0.14)	(0.24)
L. Inflation	-0.002	-0.008***	0.000	0.000	0.004	0.002	-0.038
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.03)
L. Infant mortality	-0.047	-0.034	-0.028	-0.031	-0.023	0.047	-0.032
	(0.03)	(0.06)	(0.04)	(0.04)	(0.03)	(0.04)	(0.07)
L.Exp+imp in GDP	0.006	0.003	0.004	-0.003	0.018	-0.029	0.023
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.05)
L. Agr. VA	0.003	-0.060	-0.026	-0.025	-0.109**	-0.075	-0.051
	(0.04)	(0.09)	(0.06)	(0.06)	(0.05)	(0.11)	(0.19)
Constant	17.029***	18.732***	14.754***	14.580***	15.400***		14.023**
	(5.34)	(5.92)	(4.94)	(4.79)	(4.12)		(6.31)
F-statistic	12.34	15.14	16.46	18.65	19.49	9.30	7.41
Obs.	2024	2024	2024	2024	2024	1904	538
Nr of instruments	256	46	60	74	42	52	19
ABAR(1)	0.00	0.00	0.00	0.00	0.00	0.00	0.00
ABAR(2)	0.74	0.61	0.82	0.86	0.82	0.73	0.25
Sargan p-val	0.00	0.00	0.00	0.00	0.00	0.00	0.08
Hansen p-val	1.00	0.21	0.21	0.42	0.14	0.62	0.75

Notes: Column numbers relate to the seven models listed in the text. The dependent variable is the share of government expenditures in GDP (%). *Total aid* is measured as the share of aid in GDP (%). Year and country fixed effects are included in all models, coefficients are not reported. Standard errors are clustered by country and are reported in parentheses. Two-step GMM system estimator using Windmeijer's(2005) correction is used. Sargan p-val and Hansen p-val show p-value of Sargan and Hansen's J-test on instrument validity, AB AR(1) and AB AR(2) show p-values for Arellano and Bond's (1991) test of first- and second-order serial correlation in the differenced residuals. $^{*} p < 0.1, ^{**} p < 0.05, ^{***} p < 0.01$

1. The system estimator, optimal PCA (eigenvalues of one and above are retained).
2. The system estimator, number of components set at 30.
3. The system estimator, number of components set at 50.
4. The system estimator, number of components set at 70.
5. The system estimator, optimal PCA, lagged controls strictly exogenous.
6. The difference estimator, optimal PCA (eigenvalues of one and above are retained).

Table 15: Results for the system and difference GMM estimator for 4 year averages

	[1]	[2]	[3]	[4]	[5]	[6]
Total aid	0.152	0.166	0.139	0.204*	-0.065	0.106
	(0.11)	(0.11)	(0.12)	(0.10)	(0.37)	(0.17)
L.Gov. exp. in GDP	0.773***	0.884***	0.914***	0.809***	0.835***	0.502
	(0.18)	(0.23)	(0.17)	(0.17)	(0.22)	(0.34)
GDP growth	-0.050	0.393	0.186	0.013	-0.877	-0.212
	(0.17)	(0.48)	(0.31)	(0.21)	(0.65)	(0.28)
L. Inflation	-0.001	0.000	-0.001	-0.001	-0.002	-0.002*
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
L. Infant mortality	0.016	0.064	0.038	0.027	0.009	-0.027
	(0.04)	(0.06)	(0.05)	(0.04)	(0.04)	(0.06)
L.Exp+imp in GDP	0.031	0.044	0.028	0.037	0.013	-0.024
	(0.03)	(0.04)	(0.04)	(0.03)	(0.02)	(0.05)
L. Agr. VA	-0.147	-0.239	-0.172	-0.173	0.018	-0.298
	(0.12)	(0.18)	(0.13)	(0.12)	(0.09)	(0.23)
Constant	4.930	-3.416	-1.552	2.105	8.252	
	(7.91)	(13.35)	(10.71)	(8.55)	(7.62)	
F-statistics	22.20	18.98	31.96	26.85	69.56	5.81
Obs.	474	474	474	474	474	356
Nr of instruments	56	21	35	49	17	27
ABAR(1)	0.00	0.01	0.00	0.00	0.01	0.24
ABAR(2)	0.02	0.04	0.03	0.03	0.11	0.03
Sargan p-val	0.04	0.21	0.21	0.06	0.40	0.13
Hansen p-val	0.07	0.76	0.62	0.21	0.68	0.41

Notes: Column numbers relate to the six models listed in the text. The dependent variable is the share of government expenditures in GDP (%). *Total aid* is measured as the share of aid in GDP (%). Year and country fixed effects are included in all models, coefficients are not reported. Standard errors are clustered by country and are reported in parentheses. Two-step GMM system estimator using Windmeijer's(2005) correction is used. Sargan p-val and Hansen p-val show p-value of Sargan and Hansen's J-test on instrument validity, AB AR(1) and AB AR(2) show p-values for Arellano and Bond's (1991) test of first- and second-order serial correlation in the differenced residuals. $^{*}p < 0.1,^{**}p < 0.05,^{***}p < 0.01$

The results for the four year average are presented in Table 15. The Sargan test rejects the null hypothesis of the validity of instrument set. The Arellano-Bond autocorrelation test fails to reject the null hypothesis of autocorrelation for some models. The impact of aid is always statistically insignificant, and even negative when the unlikely assumption of strict exogeneity of all lagged variables is made. Table 22 in Appendix A shows similar results for the five year average. For the three year average presented in Table 20, the optimal PCA returns positive and statistically significant impact of aid on government

Table 16: Results for the system and difference GMM estimator using factorized instruments for 4 year averages

	[1]	[2]	[3]	[4]	[5]	[6]
Total aid	0.116	0.100	0.107	0.088	-0.011	0.106
	(0.08)	(0.21)	(0.09)	(0.06)	(0.14)	(0.14)
L.Gov. exp. in GDP	0.638***	0.509*	0.651***	0.761***	0.566**	0.258
	(0.15)	(0.29)	(0.15)	(0.08)	(0.24)	(0.27)
GDP growth	0.123	0.143	0.158	0.002	0.185	-0.124
	(0.21)	(0.32)	(0.23)	(0.18)	(0.36)	(0.30)
L. Inflation	-0.001	-0.001	-0.000	-0.001*	-0.001	-0.001
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
L. Infant mortality	0.004	0.016	0.008	-0.005	-0.009	-0.062
	(0.02)	(0.06)	(0.03)	(0.02)	(0.02)	(0.05)
L.Exp+imp in GDP	-0.020	0.007	-0.017	0.020	0.018	-0.045
	(0.03)	(0.05)	(0.03)	(0.02)	(0.02)	(0.06)
L. Agr. VA	-0.191*	-0.151	-0.195*	-0.032	-0.082*	-0.107
	(0.10)	(0.16)	(0.10)	(0.05)	(0.05)	(0.27)
Constant	12.996**	12.664*	12.315**	5.442	11.350	
	(5.37)	(7.24)	(5.57)	(3.34)	(7.89)	
F-statistic	205.57	122.03	194.75	285.27	56.30	5.61
Obs.	474	474	474	474	474	356
Nr of instruments	55	37	57	77	33	35
ABAR(1)	0.01	0.16	0.01	0.00	0.10	0.62
ABAR(2)	0.04	0.09	0.04	0.04	0.05	0.09
Sargan p-val	0.07	0.05	0.07	0.02	0.02	0.01
Hansen p-val	0.66	0.29	0.64	0.28	0.66	0.17

Notes: Column numbers relate to the six models listed in the text. The dependent variable is the share of government expenditures in GDP (%). *Total aid* is measured as the share of aid in GDP (%). Year and country fixed effects are included in all models, coefficients are not reported. Standard errors are clustered by country and are reported in parentheses. Two-step GMM system estimator using Windmeijer's(2005) correction is used. Sargan p-val and Hansen p-val show p-value of Sargan and Hansen's J-test on instrument validity, AB AR(1) and AB AR(2) show p-values for Arellano and Bond's (1991) test of first- and second-order serial correlation in the differenced residuals. $^{*}p<0.1,^{**}p<0.05,^{***}p<0.01$

expenditures. However, the number of instruments is very high and the estimate of the lagged government expenditures is close to one. Also, the results are not robust to the reduction of the instrument set.[62]

Overall, none of the regressions are robust to the changes in the instrument set and none perform consistently well on overidentification tests. What

[62] The Hansen-in-difference test does not raise concerns regarding the validity of additional instruments in the system GMM estimators in this subsection. All specifications were also estimated using only the difference GMM estimator or the forward orthogonal deviations method. The results remain unstable.

is more, the estimators do not bring increased efficiency, as the results are insignificant. It is likely that the lagged levels or lagged differences of the variables of main interest are not good instruments for differences or levels respectively. Also, the estimates may be volatile because of the bias caused by the large number of instruments. There is also a possibility of a structural break in the long panel. In general, these results give strong indication that internal instruments are not able to account for the endogeneity problem and that both difference and system GMM estimators are not helpful in establishing the causal relation between aid and government expenditures.

Results based on the Chatterjee, Giuliano, Kaya (2012) dataset

Chatterjee et al. (2012) use equation (3.1) to estimate the impact of aid on government expenditures. Their dataset consists of 67 countries for the years 1972–2000, and contains around 650 observations.[63] Control variables include: GDP growth, lagged infant mortality rate, lagged agricultural value added, lagged literacy rate, lagged sum of exports and imports in GDP, lagged dependency rate, lagged total debt service in GDP, and lagged overall budget balance in GDP. Chatterjee et al. (2012) find that the coefficient of aid impact is 0.3. This result is robust to causality checks where aid is instrumented using the difference and system GMM estimators and where the instruments based on geographical and cultural proximity are used. The latter instrument is discussed in the next section.

The dynamic panel model is not used, therefore the difference and system GMM estimators are only used to account for the endogeneity of aid and other regressors. I use their specification and dataset to test the robustness of both difference and system GMM estimators.

In their paper, Chatterjee et al. (2012) do not provide information on how the difference GMM estimator is specified (the system GMM estimator is not reported). However, all specification choices are well documented in the code

[63]This dataset is considerably different from the dataset used in the previous section. It covers the additional years 1972–1979, however misses 2001–2012 (for which more complete data are available). Also, the number of countries is smaller. These differences lead to a significantly smaller number of observations. 46 countries coincide in both datasets.

files I have received from the authors:[64] aid and growth are endogenous, the lagged dependent variable and all lagged controls are predetermined, the one-step estimator is used, the instrument set is limited to one lag, forward orthogonal deviations transformation of the data is used (therefore I call it the FOD estimator and not the difference estimator), and errors are clustered at the country level. The specification remains the same for the system estimator. For both estimators, time dummies are used as regressors but are not included in the instrument set.

First, I replicate the results from Chatterjee et al. (2012). Afterwards, I test the robustness of the results. In all checks I assume that aid and growth are endogenous, the lagged dependent variable and lagged controls are predetermined, and standard errors are clustered at the recipient level. The following seven specifications are used:[65]

1. The FOD one-step estimator, the instrument set lag limited to one (as in Chatterjee et al. 2012).
2. The FOD two-step estimator with the Windmeijer correction, the instrument set collapsed and lag limited to four, time dummies treated as strictly exogenous.
3. As above, but the instrument set both collapsed and lag limited to two.
4. The FOD one-step system estimator, the instrument set lag limited to one (as in Chatterjee et al. 2012).
5. The FOD two-step system estimator with the Windmeijer correction, the instrument set collapsed and lag limited to three, time dummies treated as strictly exogenous.
6. As above, but the instrument set both collapsed and lag limited to one.
7. The system estimator, the instrument set limited by the PCA, time dummies treated as strictly exogenous.

[64] I have received the dataset and code to the working paper (Chatterjee et al. 2007) and to a significantly revised, but not final, version of the paper. I am grateful to Santanu Chatterjee and Ilker Kaya for sending the files and for answering all my questions. These results will be updated once I receive the final version of the code. According to the explanations I have received in the emails, the number of observations in the final dataset is slightly different. Moreover, in all specifications I included *L.Exp+imp in GDP* variable in the instrument set as predetermined, since it was missing in the code I have received.

[65] For the FOD one-step estimator, the model is exactly identified if the instrument set is collapsed and lags are limited to one. To avoid this, I use two lags in specification three. For the system estimator, to decrease the number of instruments below 67, the maximum number of lags used is three. However, the results remain similar for both estimators if other lags are used as instruments.

Table 17: Tests of Chatterjee et al. (2012) FOD and system GMM regressions

	[1]	[2]	[3]	[4]	[5]	[6]	[7]
Total aid	0.292***	0.313	0.246	0.139	-0.034	-0.025	0.198
	(0.08)	(0.28)	(0.24)	(0.10)	(0.26)	(0.17)	(0.33)
GDP growth	-0.001	0.001	0.000	-0.000	-0.000	-0.000	-0.000
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
L.Debt service	0.091	0.221	0.336	0.558***	0.198	0.219	-0.099
	(0.09)	(0.24)	(0.22)	(0.14)	(0.15)	(0.19)	(0.30)
L.Budget balance	-0.159	-0.218	-0.298	-0.619***	-0.315**	-0.311**	-0.234
	(0.13)	(0.22)	(0.19)	(0.12)	(0.15)	(0.14)	(0.21)
L.Infant mortality	-0.044	-0.048	-0.003	0.060*	-0.053	-0.020	-0.028
	(0.05)	(0.17)	(0.14)	(0.04)	(0.09)	(0.15)	(0.12)
L.Agr. VA	-0.192**	-0.013	0.011	-0.111	-0.145	0.143	-0.236
	(0.09)	(0.14)	(0.32)	(0.08)	(0.28)	(0.20)	(0.24)
L.Literacy rate	-0.212	0.198	0.045	0.019	-0.197	-0.185	-0.175
	(0.20)	(0.68)	(0.66)	(0.07)	(0.15)	(0.30)	(0.17)
L.Exp+imp in GDP	-0.086**	-0.040	-0.044	0.103***	0.099	0.111**	0.048
	(0.04)	(0.07)	(0.09)	(0.02)	(0.06)	(0.05)	(0.06)
L.Dependency, 65+	-0.195	-1.144	-1.009	0.903***	1.202	1.656	1.483*
	(0.93)	(4.40)	(3.11)	(0.26)	(0.74)	(1.00)	(0.87)
Constant				-0.766	25.924	8.775	28.621
				(6.73)	(20.46)	(29.99)	(20.47)
R-sqr							
F-statistic	33.52	6.23	13.98	65.38	31.18	8.90	5.17
Obs.	516	516	516	574	574	574	574
Nr of instruments	229	63	45	442	64	46	365
ABAR(1)	0.02	0.07	0.03	0.01	0.09	0.07	0.02
ABAR(2)	0.21	0.49	0.19	0.06	0.34	0.14	0.27
Sargan p-val	0.68	0.25	0.04	0.00	0.00	0.00	0.00
Hansen p-val	1.00	0.90	0.59	1.00	0.76	0.58	1.00

Notes: The dependent variable is the share of government expenditures in GDP (%). *Total aid* is measured as the share of aid in GDP (%). Year and country fixed effects are included in all models, coefficients not reported. Standard errors clustered by country, reported in parentheses. Sargan p-val and Hansen p-val show p-value of Sargan and Hansen's J-test on instrument validity, AB AR(1) and AB AR(2) show p-values for Arellano and Bond's (1991) test of first- and second-order serial correlation. $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$

Table 17 presents the results. The first specification (and the fourth for the unreported system estimator) is taken from the code files I have received from Chatterjee et al. (2012) and should replicate their results from p.16, Table 1, column 7. The coefficient of aid impact equals 0.292 (column [1]) and is smaller than the 0.370 reported in the original paper (most likely due to slight differences in the dataset). The p-value of the Hansen statistic is identical, and the p-values of the Arellano-Bond autocorrelation tests are very close. The null hypothesis of first-order autocorrelation - ABAR(1) - is rejected.[66] The instrument count, unreported in the original paper, equals 229: substantially more than the number of units (67 countries). Therefore, the model is overfitted and the results converge towards the fixed effects results.

In other specifications (column [2], [3], [5], [6]) I decrease the number of instruments by collapsing the instrument set and limiting the lags. Moreover, I include time dummies as strictly exogenous variables, as all regressors should be included in the instrument matrix (Roodman, 2009a). Time dummies should reduce the risk of cross-country correlation. For both specifications, the number of instruments falls below the rule-of-thumb threshold value of 67, which greatly reduces the risk of overfitting. Moreover, I use the preferred two-step estimator with the Windmeijer's (2005) correction. Both, for the FOD estimator and for the system estimator Arellano-Bond autocorrelation tests indicate problems. The Sargan test raises concerns regarding the instrument validity (columns [3]-[7]). The impact of aid is insignificant, and for the system estimator the coefficients are negative, which is highly unlikely. The results are similar if the first-difference transformation is used instead of the orthogonal transformation and when one-step estimator is used.[67]

Overall, the forward orthogonal deviations and system GMM estimates do not pass the sensitivity analysis. Statistically significant results for DAC aid's impact on government expenditures in Chatterjee et al. (2012) are biased because of overfitting.

3.3 Instrument set based on cultural and geographical proximity

Instruments in a panel data analysis must vary over time, which restricts the number of possible exogenous and valid variables that could account for

[66]For first-difference transformation first order autocorrelation (rejecting the null hypothesis of ABAR(1)) is expected. However, in the orthogonal transformation the average of all available future observations is subtracted. Therefore, for correctly specified model there should be no first- or second-order autocorrelation.

[67]Available upon request.

endogeneity. Tavares (2003) and Larraín and Tavares (2004) propose a set of instruments that relies on both geographical and cultural distance between the donor and the recipient (or between trade partners), and varies over time. They use these instruments to estimate the impact of aid on corruption and the impact of trade openness on corruption respectively.

A similar approach has been tried by Chatterjee et al. (2012) to investigate the impact of aid on government expenditures. The inverse of the bilateral distance and a contiguity are used as measures of geographical proximity, and the same major religion and a common official language are used as measures of cultural proximity. They find evidence of a statistically significant impact of DAC aid on government expenditures: government expenditures increase by 37% of aid. In this chapter, I follow this approach and try to validate their results using the most recent and more complete data, and a bigger set of both donor and recipient countries. The instrument set is constructed as follows:

1. 28 out of 29 DAC donors are included in the sample: Australia, Austria, Belgium, Canada, Czech Republic, Denmark, Finland, France, Germany, Greece, Iceland, Ireland, Italy, Japan, Luxembourg, the Netherlands, New Zealand, Norway, Poland, Portugal, South Korea, Slovakia, Slovenia, Spain, Sweden, Switzerland, the United Kingdom, the United States of America. The European Union is the only DAC donor left out.[68]
2. For each donor-recipient pair in the sample four (time-invariant) variables that indicate both geographic and cultural closeness between each donor and each recipient are computed. The variables are: the inverse of bilateral distance in kilometers, a dummy variable equal to one if the country pair has a common land border, a measure of religious affinity, and a dummy equal to one if the country pair shares an official language. The French research center in international economics CEPII provides comprehensive datasets containing geographical and cultural country characteristics, including bilateral data.[69] The *Language* dataset from Melitz and Toubal (2014) is used as it contains all of the above mentioned variables.

[68]It is listed as a bilateral donor by the DAC. However, data required to construct the instrument are not available.

[69]Databases can be downloaded from http://www.cepii.fr/. This website contains also detailed information on the methodology of data construction.

3. The constant US dollar value of total aid outflows from each donor is taken from the Table DAC1a from the Development Assistance Committee of the OECD.[70] Outflows from each donor are multiplied by the variables constructed in the previous step, creating time-varying variables. As a result, for each recipient and each of the four variables there are 28 values (one for each donor). The sum of these 28 values in each of the four categories (distance, contiguity, religion, and language) constitutes the four instruments for the share of aid in GDP of each country in the sample.

   Each country in the sample will have four time-varying exogenous variables that will serve as instruments for its share of aid in GDP, defined as:

- $AID\_DIST_i = \sum_{j=1}^{28} (Inverse\ \ of\ \ Bilateral\ \ Distance_{i,j}) * Aid_j$
- $AID\_CONT_i = \sum_{j=1}^{28} Contiguous_{i,j} * Aid_j$
- $AID\_REL_i = \sum_{j=1}^{28} Religion_{i,j} * Aid_j$
- $AID\_LANG_i = \sum_{j=1}^{28} Language_{i,j} * Aid_j$

where j refers to the 28 DAC donor countries, and i to the recipient countries. The instrument set relies on the assumption that countries that have closer links to the donor receive more aid compared to the countries that are more culturally or geographically separated. The motivation for this instrument comes from the aid allocation literature. Cultural and historic ties have been found to explain a large part of the variation in aid allocation (Alesina and Dollar, 2000). For example, the main recipients of foreign aid from the United Kingdom and France are their former colonies (Boone, 1996), i.e. countries that usually share the same language and religion. Therefore, an increase in total aid outflows from a DAC country should primarily affect countries that have these ties. In the dataset provided by CEPII there is an additional variable that describes the cultural and historical ties between recipients and donors: colonial relation in 1945 (AID_COL_i).[71] Following the steps 1 to 3, I add this variable to the instrument set to check the robustness of the results. Colonial

[70] Available at http://stats.oecd.org/, under Development → Other.

[71] Two countries account for the majority of colonies in the sample: France and the United Kingdom. Australia, Belgium jointly with Luxembourg, Japan, the Netherlands, Portugal and the US are the other donors that have colonial ties in the sample.

history is highly correlated with the common language dummy, but not with the religious proxy. It is expected that the only channel of impact of these variables on the share of total government expenditures in GDP is through the share of aid in GDP. These instruments vary over time, therefore they are appropriate for the fixed effects analysis.

Table 18: Regressions for the instrument set based on cultural and geographical proximity

	[1]	[2]	[3]	[4]	[5]	[6]
Total aid	0.481*	-0.015	0.635*	-0.182	0.584**	0.117
	(0.27)	(0.20)	(0.35)	(0.34)	(0.26)	(0.18)
L.Gov. exp. in GDP		0.635***		0.665***		0.612***
		(0.05)		(0.07)		(0.05)
GDP growth	-0.015	0.001	-0.014	0.000	-0.014	0.002
	(0.05)	(0.02)	(0.06)	(0.03)	(0.06)	(0.03)
L. Inflation	0.004	0.000	0.005*	-0.000	0.004*	0.001
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
L. Infant mortality	-0.019	-0.005	-0.025	0.003	-0.023	-0.010
	(0.02)	(0.01)	(0.02)	(0.02)	(0.02)	(0.01)
L.Exp+imp in GDP	-0.006	0.009	-0.013	0.016	-0.010	0.004
	(0.01)	(0.01)	(0.02)	(0.02)	(0.01)	(0.01)
L. Agr. VA	-0.125**	-0.038	-0.105	-0.052	-0.112**	-0.026
	(0.06)	(0.04)	(0.06)	(0.04)	(0.05)	(0.03)
Constant	26.734***	11.182***	25.436***	11.457***	25.867***	10.964***
	(3.34)	(1.80)	(3.75)	(2.03)	(3.24)	(1.71)
N	1880	1797	1880	1797	1880	1797
R square (centered)	0.85	0.91	0.82	0.90	0.83	0.92
Weak id. statistic	4.38	4.19	3.03	3.02	3.59	3.49
Anderson-Rubin p-val	0.00	0.16	0.00	0.10	0.00	0.12
Endogeneity p-val	0.22	0.63	0.10	0.67	0.10	0.89

Notes: The dependent variable is the share of government expenditures in GDP (%). *Total aid* is measured as the share of aid in GDP (%). Year and country fixed effects are included in all models, coefficients not presented. AID_DIST_i, AID_CONT_i, AID_REL_i used as instruments for aid in all columns, AID_LANG_i in columns [1],[2],[5], and [6], and AID_COL_i in columns [3]–[6]. Robust standard errors, reported in parentheses. $^{*} p < 0.1,^{**} p < 0.05,^{***} p < 0.01$

The dataset from chapter two is used.[72] Both equations (3.1) and (3.2) are estimated. In the first stage, the share of aid in GDP of each developing country is regressed on the instrument set, fixed effects and other control variables.

[72]Due to missing data, the following countries are excluded from the analysis: Botswana, Ethiopia, Lesotho, Maldives, Mongolia, Montenegro, Myanmar, Namibia, Samoa, Serbia, Swaziland, Timor-Leste. The sample consist of 106 countries.

Three instrument sets are used. The first consist of the initial 4 instruments that are interacted with aid: bilateral distance, contiguity, common language, and the index of religious proximity. In the second instrument set, common language is replaced by the highly correlated colonial past. In the third instrument set, all five instruments are included. If the reasons behind the construction of the instrument set are correct, all the instruments should have a positive sign in the first stage regression. A decrease in the distance (so an increase in the inverse distance), sharing common border, religion, or language should increase the amount of aid. Table 23 in Appendix B presents the first stage regressions.[73]

The predicted values of the dependent variable from the first stage is then used in the second stage to explain the share of government expenditures in GDP. The results are presented in Table 18. The Anderson-Rubin test (row *Anderson-Rubin p-val*) of the significance of endogenous regressors (robust to the presence of weak instruments) rejects the null hypothesis and indicates that the instruments are relevant for models that do not include a lagged dependent variable. The endogeneity test (*Endogeneity p-val*) has the null hypothesis that the specified endogenous regressors can be treated as exogenous. For all models, the null hypothesis cannot be rejected at 5% level, which indicates that aid is exogenous to government expenditures and that the fixed effects estimates presented in chapter two may be unbiased. Weak identification statistics (*Weak id. statistic*) gives the Wald F-value for the Kleibergen-Paap *rk* statistic. The values are below the critical values provided by Stock and Yogo (2005), which means that I cannot rule out a weak-instrument problem. That suggests that the model is not adequately identified, although Baum et al. (2007) point out that these statistics require i.i.d. errors and thus results should be approached with caution. They suggest looking also at the F-statistic of the excluded instruments in the first stage and follow the well known "rule

[73] For my sample, contiguity has a negative statistically significant impact in the first stage. However, there are only few pairs: Greece is bordering Albania, Macedonia and Turkey, and Poland and Slovakia are bordering Ukraine. While a significant share of total aid from Poland and Slovakia is going to Ukraine (however, it is a low amount in absolute term), the outflows from Greece to Macedonia and Turkey are low, likely due to political tensions between these countries. The distance variable is positive, however insignificant. The common language variable is always positive and statistically significant, and the religion proxy only when the common language variable is replaced by colonial history.

of thumb" of Staiger and Stock (1997). F-statistic values below 10, as it is the case for my regressions (see Table 23 in Appendix B), indicate that weak identification is a potential problem. Summing up, the instruments do not pass all the canonical tests which suggests that the results may be biased due to weak identification.

These caveats should be kept in mind when interpreting the results. I find that aid increases government expenditures by 48.1% to 63.5% depending on the instrument set used, however the standard errors are large. In the study that uses fixed effects and does not instrument aid (chapter two), I find that government expenditures increase by around 15% of aid for the same equation (3.1). The IV estimate is considerably higher and is close to the long-term estimates when the averaged data and dynamic models are used. For the models with the lagged dependent variable total aid is insignificant, and in columns [2] and [4] has a negative sign that is highly unlikely. Only in column [6] the estimate is close to the estimates from chapter two, however here it is not statistically significant.

Chatterjee et al. (2012) finds statistically stronger results.[74] The sign of instruments in the first stage is, as expected, positive and three of them (distance, religion and language) are statistically significant for DAC data, as well as distance and religion for the CRS data. In the footnote the authors note that this specification passes the canonical tests for identification and instrumental relevance. Chatterjee et al. (2012) find that the impact of the instrumented aid (taken from the DAC database) is statistically significant, positive and on average 20.1% of it increases government expenditures, while 79.9% is fungible. These results are close to the fixed effects estimates (25.9% of aid increases government expenditures), and to the Arellano-Bond GMM estimates discussed in the previous section (37%). The results for the two aid variables taken from the Creditor Reporting System are statistically insignificant.

3.4 Conclusions

This chapter investigates two methods that have been employed in the fungibility literature to estimate the impact of aid on total government expenditures.

[74] As previously discussed, Chatterjee et al. (2012) use different dataset. Moreover, they use a binary variable for religious proximity.

First, I discuss the difference and system GMM estimators and point out the potential weak spots of these methods of estimation. Then I apply both estimators to the dataset used in chapter two. I report all the specification choices and test the robustness of the results to the changes in the instrument set. The results are volatile and small changes in the specification lead to substantial changes in the results. I show that unless these estimators are applied with great care, they can easily generate misleading results. This may lead to a confirmation bias, a situation where the researcher stops looking for other specifications or other methods because the expected results (e.g. ones that confirm the hypothesis or are close to the results found with the fixed effects) have been found. I also use the dataset and specification of Chatterjee et al. (2012) and show that the results are biased because of overfitting and that these GMM estimators are not able to establish a causal relationship between aid and government expenditures. For both datasets, the mostly insignificant results should not be interpreted as lack of impact of aid on government expenditures, but as a sign of the weakness of these methods in this particular application. Long panels increase the number of potential instruments and increase the probability of overfitting. Moreover, internal instruments seem to be weak in these applications.

Second, I use an instrument set based on cultural and geographical proximity to estimate the impact of aid on government expenditures. For the dynamic specification, the coefficient of aid is close to zero and statistically significant. When lagged government expenditures are not included, the government expenditures increase by between 48% and 64% of aid. The instrument is weak according to various specification tests, which suggests that the results may be biased.

3.5 Appendix

Appendix A: Difference and system GMM results for the dataset from chapter two averaged over 3 and 5 year

Table 19: Results for the system and difference GMM estimator for 3 year averages

	[1]	[2]	[3]	[4]	[5]	[6]
Total aid	0.061	0.034	0.084	0.110	0.059	0.082
	(0.09)	(0.14)	(0.17)	(0.12)	(0.13)	(0.15)
L.Gov. exp. in GDP	0.659***	0.905***	0.730***	0.760***	0.881***	0.436**
	(0.10)	(0.18)	(0.17)	(0.16)	(0.23)	(0.22)
GDP growth	0.029	0.629	0.342	0.253	-0.222	0.458
	(0.17)	(0.43)	(0.39)	(0.28)	(0.63)	(0.34)
L. Inflation	-0.005	-0.003	-0.002	-0.002	0.002	-0.004
	(0.01)	(0.01)	(0.01)	(0.01)	(0.00)	(0.01)
L. Infant mortality	0.015	0.101*	0.028	0.014	-0.003	0.001
	(0.03)	(0.06)	(0.04)	(0.03)	(0.01)	(0.04)
L.Exp+imp in GDP	0.030	0.041	0.019	0.020	0.003	-0.040
	(0.02)	(0.04)	(0.04)	(0.03)	(0.01)	(0.03)
L. Agr. VA	-0.109	-0.257	-0.179	-0.147	-0.021	-0.288
	(0.09)	(0.16)	(0.12)	(0.10)	(0.05)	(0.18)
Constant	8.619**	-2.581	7.162	6.305	5.068	
	(4.22)	(8.20)	(6.62)	(5.85)	(7.95)	
F-statistic	17.02	10.97	12.42	16.47	38.56	8.27
Obs.	650	650	650	650	650	531
Nr of instruments	80	24	38	52	20	30
ABAR(1)	0.00	0.00	0.01	0.00	0.01	0.05
ABAR(2)	0.21	0.11	0.14	0.19	0.17	0.15
Sargan p-val	0.01	0.43	0.06	0.01	0.01	0.18
Hansen p-val	0.72	0.91	0.62	0.72	0.30	0.72

Notes: Column numbers relate to the six models listed in the text. The dependent variable is the share of government expenditures in GDP (%). *Total aid* is measured as the share of aid in GDP (%). Year and country fixed effects are included in all models, coefficients are not reported. Standard errors are clustered by country and are reported in parentheses. Two-step GMM system estimator using Windmeijer's(2005) correction is used. Sargan p-val and Hansen p-val show p-value of Sargan and Hansen's J-test on instrument validity, AB AR(1) and AB AR(2) show p-values for Arellano and Bond's (1991) test of first- and second-order serial correlation in the differenced residuals. $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$

Table 20: Results for the system and difference GMM estimator using factorized instruments for 3 year averages

	[1]	[2]	[3]	[4]	[5]	[6]
Total aid	0.214***	0.038	0.195**	0.194***	0.117	0.178*
	(0.07)	(0.11)	(0.08)	(0.07)	(0.08)	(0.09)
L.Gov. exp. in GDP	0.898***	0.474	0.790***	0.897***	0.536***	0.295
	(0.06)	(0.30)	(0.12)	(0.08)	(0.13)	(0.20)
GDP growth	-0.139	-0.273	-0.177	-0.157	-0.208	-0.020
	(0.18)	(0.32)	(0.19)	(0.16)	(0.22)	(0.19)
L. Inflation	-0.002	-0.021*	-0.001	0.000	0.003	0.002
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
L. Infant mortality	-0.021	-0.011	-0.034*	-0.039**	-0.018	-0.031
	(0.02)	(0.04)	(0.02)	(0.02)	(0.02)	(0.03)
L.Exp+imp in GDP	0.016	-0.034	-0.010	-0.009	0.023	-0.007
	(0.02)	(0.05)	(0.04)	(0.03)	(0.02)	(0.04)
L. Agr. VA	-0.089	-0.015	-0.106	-0.072	-0.065*	-0.091
	(0.06)	(0.25)	(0.13)	(0.06)	(0.04)	(0.20)
Constant	4.301	20.511	10.635*	7.109*	14.066***	
	(2.81)	(12.88)	(5.47)	(3.65)	(3.86)	
F-statistic	34.89	7.12	46.68	25.93	21.48	4.57
Obs.	650	650	650	650	650	531
Nr of instruments	91	40	60	80	57	66
ABAR(1)	0.00	0.26	0.00	0.00	0.01	0.17
ABAR(2)	0.23	0.61	0.31	0.23	0.24	0.33
Sargan p-val	0.00	0.74	0.01	0.00	0.00	0.00
Hansen p-val	0.48	0.45	0.51	0.31	0.35	0.24

Notes: Column numbers relate to the six models listed in the text. The dependent variable is the share of government expenditures in GDP (%). *Total aid* is measured as the share of aid in GDP (%). Year and country fixed effects are included in all models, coefficients are not reported. Standard errors are clustered by country and are reported in parentheses. Two-step GMM system estimator using Windmeijer's(2005) correction is used. Sargan p-val and Hansen p-val show p-value of Sargan and Hansen's J-test on instrument validity, AB AR(1) and AB AR(2) show p-values for Arellano and Bond's (1991) test of first- and second-order serial correlation in the differenced residuals. $^{*} p < 0.1, ^{**} p < 0.05, ^{***} p < 0.01$

Table 21: Results for the system and difference GMM estimator for 5 year averages

	[1]	[2]	[3]	[4]	[5]	[6]
Total aid	0.220**	0.009	0.248**	0.220**	-0.043	0.331
	(0.10)	(0.21)	(0.11)	(0.10)	(0.49)	(0.30)
L.Gov. exp. in GDP	0.740***	1.078***	0.815***	0.740***	0.896**	0.251
	(0.16)	(0.23)	(0.22)	(0.16)	(0.41)	(0.39)
GDP growth	0.258	0.745	0.465	0.258	-0.483	0.001
	(0.25)	(0.52)	(0.33)	(0.25)	(0.98)	(0.46)
L. Inflation	-0.001	-0.001	-0.001	-0.001	-0.002	-0.001
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
L. Infant mortality	0.047	0.111	0.055	0.047	-0.001	-0.061
	(0.04)	(0.10)	(0.05)	(0.04)	(0.05)	(0.09)
L.Exp+imp in GDP	0.070**	0.035	0.049	0.070**	0.008	-0.043
	(0.03)	(0.05)	(0.04)	(0.03)	(0.03)	(0.07)
L. Agr. VA	-0.240**	-0.215	-0.279**	-0.240**	0.028	-0.272
	(0.11)	(0.16)	(0.12)	(0.11)	(0.08)	(0.25)
Constant	3.487	-6.661	2.557	3.487	6.373	
	(5.95)	(9.21)	(7.14)	(5.95)	(13.55)	
F-statistic	26.55	19.14	27.10	26.55	32.33	5.65
Obs.	364	364	364	364	364	247
Nr of instruments	40	19	33	40	15	25
ABAR(1)	0.00	0.00	0.01	0.00	0.15	0.75
ABAR(2)	0.92	0.82	0.79	0.92	0.78	0.62
Sargan p-val	0.00	0.07	0.00	0.00	0.01	0.03
Hansen p-val	0.49	0.58	0.43	0.49	0.09	0.21

Notes: Column numbers relate to the six models listed in the text. The dependent variable is the share of government expenditures in GDP (%). *Total aid* is measured as the share of aid in GDP (%). Year and country fixed effects are included in all models, coefficients are not reported. Standard errors are clustered by country and are reported in parentheses. Two-step GMM system estimator using Windmeijer's(2005) correction is used. Sargan p-val and Hansen p-val show p-value of Sargan and Hansen's J-test on instrument validity, AB AR(1) and AB AR(2) show p-values for Arellano and Bond's (1991) test of first- and second-order serial correlation in the differenced residuals. $^{*}p<0.1,^{**}p<0.05,^{***}p<0.01$

Table 22: Results for the system and difference GMM estimator using factorized instruments for 5 year averages

	[1]	[2]	[3]	[4]	[5]	[6]
Total aid	0.067	0.088	0.092	0.210***	0.029	0.085
	(0.19)	(0.20)	(0.11)	(0.08)	(0.10)	(0.25)
L.Gov. exp. in GDP	0.580***	0.548**	0.645***	0.735***	0.332*	0.482
	(0.21)	(0.23)	(0.14)	(0.12)	(0.17)	(0.34)
GDP growth	0.175	0.188	0.315	0.069	-0.566	-0.135
	(0.22)	(0.23)	(0.22)	(0.15)	(0.42)	(0.34)
L. Inflation	-0.001**	-0.001**	-0.001	-0.001	- 0.002***	-0.002
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
L. Infant mortality	0.030	0.027	-0.024	-0.036*	-0.021	-0.057
	(0.07)	(0.06)	(0.03)	(0.02)	(0.02)	(0.08)
L.Exp+imp in GDP	-0.020	-0.019	-0.030	-0.003	0.028*	-0.066
	(0.05)	(0.05)	(0.02)	(0.02)	(0.02)	(0.10)
L. Agr. VA	-0.166	-0.186	-0.096	-0.068	-0.062	-0.328
	(0.13)	(0.12)	(0.07)	(0.06)	(0.05)	(0.62)
Constant	15.998***	17.136***	15.358***	11.181***	21.986***	
	(5.49)	(5.51)	(4.92)	(3.89)	(4.96)	
F-statistic	25.14	28.25	21.70	25.18	28.15	4.51
Obs.	364	364	364	364	364	247
Nr of instruments	34	35	55	75	22	20
ABAR(1)	0.13	0.17	0.03	0.00	0.61	0.53
ABAR(2)	0.93	0.91	0.89	0.86	0.81	0.78
Sargan p-val	0.02	0.01	0.00	0.00	0.01	0.59
Hansen p-val	0.12	0.17	0.31	0.25	0.62	0.35

Notes: Column numbers relate to the six models listed in the text. The dependent variable is the share of government expenditures in GDP (%). *Total aid* is measured as the share of aid in GDP (%). Year and country fixed effects are included in all models, coefficients are not reported. Standard errors are clustered by country and are reported in parentheses. Two-step GMM system estimator using Windmeijer's(2005) correction is used. Sargan p-val and Hansen p-val show p-value of Sargan and Hansen's J-test on instrument validity, AB AR(1) and AB AR(2) show p-values for Arellano and Bond's (1991) test of first- and second-order serial correlation in the differenced residuals. $^{*}p < 0.1,^{**}p < 0.05,^{***}p < 0.01$

Appendix B: First stage results

Table 23: First stage results

	[1]	[2]	[3]	[4]	[5]	[6]
AID_CONTi	-4.345*	-5.812**	-4.727*	-6.265**	-4.522*	-6.011**
	(2.62)	(2.66)	(2.67)	(2.75)	(2.61)	(2.68)
AID_DISTi	65.462	21.218	53.730	9.299	40.533	-5.541
	(45.37)	(44.74)	(47.07)	(47.07)	(47.78)	(48.23)
AID_RELi	0.051	0.018	0.108***	0.078*	0.027	-0.009
	(0.05)	(0.05)	(0.04)	(0.04)	(0.06)	(0.06)
AID_LANGi	0.067**	0.070***			0.105**	0.112**
	(0.03)	(0.03)			(0.05)	(0.05)
AID_COLi			0.018	0.015	-0.175	-0.187
			(0.06)	(0.06)	(0.13)	(0.13)
GDP growth	-0.009	-0.004	-0.013	-0.008	-0.010	-0.005
	(0.08)	(0.08)	(0.08)	(0.08)	(0.08)	(0.08)
L. Inflation	-0.006*	-0.004	-0.005	-0.004	-0.005*	-0.004
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
L. Infant mortality	0.028	0.037	0.029	0.039	0.031	0.040
	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)
L.Exp+imp in GDP	0.047***	0.041***	0.046***	0.040***	0.045***	0.039***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
L. Agr. VA	-0.128**	-0.086	-0.133**	-0.092*	-0.133**	-0.091
	(0.05)	(0.05)	(0.06)	(0.06)	(0.05)	(0.06)
L.Gov. exp. in GDP		0.172***		0.174***		0.170***
		(0.04)		(0.04)		(0.04)
Constant	3.610	-2.059	5.755*	0.161	4.755	-0.813
	(3.22)	(3.17)	(3.37)	(3.23)	(3.33)	(3.21)
N	1880	1797	1880	1797	1880	1797
R-square (centered)	0.74	0.75	0.74	0.75	0.74	0.75
F-test	4.38	4.19	3.03	3.02	3.59	3.49

Notes: The dependent variable is the share of aid in GDP (%). Year and country fixed effects are included in all models, coefficients not presented. F-test presents the value of Angrist-Pischke multivariate F test of excluded instruments. $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$

CHAPTER 4

The Causal Links Between Aid and Government Expenditures

The most recent literature on aid effectiveness finds a positive effect of aid on growth. To the extent that aid goes through the budget, this effect reflects either an aid-financed increase in government expenditures (quantity effect) or an improvement in the use of government resources as a result of donor involvement and lower taxes (quality effect). This study investigates the causal link between on-budget aid and government expenditures using a large cross-country panel dataset for 53 countries and recent methodology to test Granger causality in heterogeneous panels. Most donor countries are found not to change aid in response to changes in government expenditures, and recipient governments react to aid by changing the way they use their own resources rather than by increasing spending. Contrary to conventional wisdom, there is little support for a quantity effect: aid Granger-causes government expenditures in only eight countries. This suggests that aid substitutes domestic government revenue or borrowing and that aid is largely effective through the quality effect.

I am especially grateful to Chris Elbers and Jan Willem Gunning for many useful comments and insights during numerous talks and to Eric Bartelsman, Bas van der Klaauw, Maarten Lindeboom, Remco Oostendorp, Tymon Słoczyński, Lisette Swart, Melinda Vigh, and seminar participants at Tinbergen Institute, VU University, RGS Conference, Warsaw Economic Seminar, and Izmir University for useful suggestions. The usual disclaimers apply.

4.1 Introduction

A large body of literature on aid effectiveness has tried to assess the impact of total aid on growth, poverty reduction, and on living standards in developing countries.[75] Recently, the methodology and robustness of many of the best known aid effectiveness studies have been criticized (see, for example, Tarp 2006; Roodman 2007; Arndt et al. 2010). Most recent studies find a positive effect of total aid on growth and poverty reduction (see Arndt et al. 2010, 2011, 2014; Alvi and Senbeta 2012; Juselius et al. 2014; Mekasha and Tarp 2013; Galiani et al. 2014).[76] Through what channels this impact is achieved is less clear (Arndt et al., 2011; Bourguignon and Sundberg, 2007). When aid is channeled through the budget it can either increase government expenditures, or it can substitute other sources of government revenues or borrowing. In the former case, the positive impact of aid on growth is likely to reflect the aid-financed increase in government spending (quantity effect). In the latter case, aid effectiveness reflects changes in the structure and quality of government spending or the impact of lower taxes (quality effect).[77] To understand aid effectiveness it is therefore crucial to investigate the link between on-budget aid and government expenditures and the direction of causality.

[75] This chapter uses standard terminology from the aid effectiveness literature. All terms and concepts are understood as defined by the OECD (the glossary can be found at: http://stats.oecd.org/glossary/) and the World Bank (http://www.worldbank.org/depweb/english/modules/glossary.html#m). However, since this study investigates the importance of the government channel, it focuses only on on-budget aid (aid that is recorded in the budget). Therefore, aid and on-budget aid are used interchangeably. To make the distinction clear when it is needed, aid that includes both off- and on-budget components is called total aid.

[76] Nevertheless, the debate is still ongoing. For example, Nowak-Lehmann et al. (2012) exploit the time-series properties of the data and find insignificant impact of aid on per capita income. Their approach was criticized by Lof et al. (2014a), which was followed by a comment of Herzer et al. (2014), and a rejoinder by Lof et al. (2014b).

[77] The quality effect is used to describe the potential impact of on-budget aid on development indicators when aid is (partly or fully) fungible at the aggregate level. This means that aid does not increase government expenditures by its full amount and aid substitutes domestic revenues or borrowing. If the structure of government expenditures does not change, it means that on-budget aid is indirectly influencing growth by enabling tax reductions or decreasing borrowing. However, it is also likely that the inflow of aid changes the structure of government expenditures - e.g. more money is spent in sectors prioritized by the donors, such as education and health. Moreover, donor involvement may improve legislation in recipient country or improve the efficiency of the distribution of total government expenditures. Through all of these channels, on-budget aid can influence growth and other development indicators, even if government expenditures do not increase (so even if there is no quantity effect).

The impact of aid on government expenditures is investigated in the fungibility literature and in fiscal response studies (for a review see McGillivray and Morrissey 2004; Morrissey 2014). These literatures suffer from the same methodological weaknesses as the earlier aid-growth regressions. A potential problem is the endogeneity of aid in regressions which has been ignored in many studies that investigate the impact of foreign aid on government expenditures. Chapter two summarizes and discusses three reasons to expect endogeneity. Firstly, aid is received in packages designed for a few years and often the recipient government decides on the timing of disbursement (McGillivray and Morrissey, 2000). As a result, aid inflow may depend on the budgetary situation of recipient country's government, which leads to reverse causality. Secondly, donors may increase the amounts of aid in countries that fail to provide merit goods. Finally, some countries, possibly those with better institutions, may be allowed by the donor countries to treat part of their aid as fungible.

Since it is difficult to find good instruments for aid, I do not use an IV-approach to establish causality between government expenditures and on-budget aid.[78] Instead, I use the Granger non-causality test for heterogeneous panel data as discussed by Hurlin and Venet (2001) and Hurlin and Dumitrescu (2012), and applied by Erdil and Yetkiner (2009) and Hood et al. (2008).[79] This method can account for two types of heterogeneity (Hurlin and Venet, 2001): differences between countries in levels and slopes. The latter type of heterogeneity is the more crucial one, since if it is ignored, the general conclusion regarding the causality relationship in the sample may be biased. Moreover, the fungibility literature has focused on the short-term impact of aid on government expenditures (see chapter two). The Granger causality test in this chapter establishes the long-term link between aid and government expenditures.

I also investigate whether there are differences in causality patterns between countries. Granger causality will not be detected if aid is fungible, that

[78] Chapter three discusses two instrumental regression strategies using a similar dataset.

[79] Note that the Granger test can also give some indications regarding exogeneity. Granger causality is necessary (but not sufficient) for strong exogeneity. At the same time, it is neither necessary nor sufficient to establish weak exogeneity. However, as Maddala (2001) and Gujarati (2007) point out, the Granger test is considered to be a useful tool for description of time-series and panel data.

is, if the government is able to react to the changes in aid by changing the way it uses its own resources, so that the recipient country's level of government expenditures increases by less than the received amount of aid. Since it is easier to make relatively small amounts of aid fungible, it is expected that on-budget aid will be Granger-causing government expenditures among LDCs.[80] As a fraction of GDP, the Least Developed Countries (LDCs) receive more than ten times as much aid as the Lower and Middle Income Countries (LMICs). The second reason for endogeneity, that donors may increase the amounts of aid in countries that fail to provide merit goods, also suggests that government expenditures should be Granger-causing aid more often among the Least Developed Countries.[81]

To my knowledge, there is no cross-country study testing the Granger causality in the link between aid and government expenditures. At the country level, Osei et al. (2005) investigate the causal relation between aid and government expenditures in Ghana. They find that foreign aid is taken as given, that it is weakly exogenous, and that it Granger-causes government expenditures. Tax revenue, government spending, and borrowing, on the other hand, are adjusted in response to the imbalance in the fiscal situation. In the related literature on aid effectiveness, Arvin and Barillas (2002) test the Granger causality link between aid and the level of poverty, controlling for the level of democratization. They find substantial differences between countries at different levels of development and they comment that there is a real possibility in some cases that the disbursement of aid and poverty reduction are not linked positively. Roodman (2008) shows that growth is Granger-causing aid.

The aim of this chapter is to provide evidence on causality in the Granger sense between on-budget aid and government expenditures. Recent econometric techniques are used on the largest possible dataset. For most countries no causality between aid and government expenditures is found, which means that governments react to aid by changing the way they use their own resources

[80] It is not assumed that aid is fully fungible, but, as argued by Marć (2012), simple dynamic models suggest that an optimizing government would divide the increase in income to keep shares of private consumption and government spending constant.

[81] It is also rather probable that both links may not be detected for the Lower and Middle Income Countries due to the fact that aid constitutes only a very small share of the recipient countries' GDP in this group.

and that aid is fungible. Also, government expenditures are not Granger-causing aid, which may mean that donors are not reacting immediately to changes in recipient countries and that the commitment packages have largely predetermined disbursement timing (contrary to what McGillivray and Morrissey 2000 suggest).

The chapter is organized as follows. Section two describes the methodology. Section three focuses on data and sample characteristics. Empirical results are presented and discussed in section four. The last section concludes.

4.2 Methodology

The idea of Granger causality as described in Granger (1969, 1980) is based on the principle that a cause cannot come after its effect. When variable x is affecting a variable y, addition of the past values of the former variable should increase the precision of prediction of the latter variable (Lutkepohl, 2009). Kirchgassner and Wolters (2007) define it formally: for a weakly stationary time series (where the first and second moments are constant), an information set I_t, $\overline{x}_t$ being a set of all current and past values of x, and variance of the forecast error denoted by σ^2, Granger causality between x and y is defined as:

- *Granger causality:* x is Granger causal to y if and only if the application of an optimal linear prediction function leads to

$$\sigma^2(y_{t+1}|I_t, \overline{x}_t) < \sigma^2(y_{t+1}|I_t) \tag{4.1}$$

  In that sense, the inclusion of x leads to better prediction of y (as it reduces the forecast errors).

- *Instantaneous Granger causality:* the application of an optimal linear prediction function leads to:

$$\sigma^2(y_{t+1}|I_t, x_{t+1}) < \sigma^2(y_{t+1}|I_t) \tag{4.2}$$

  In that sense, the future value of y can be predicted better if the future value of x is used in addition to the current and past values of x.

- *Feedback:* There is a feedback between x and y if x is causal to y and y is causal to x.

In applications, regression techniques and Wald statistics are used to test Granger causality. While Granger causality has mostly been tested in time series, during the last two decades there has been a substantial development of both methods and applications of Granger causality in panel data settings, accounting for heterogeneity of panels (see Hurlin and Venet, 2001; Hurlin and Dumitrescu, 2012). Panel data have certain advantages over time series: both cross-sectional and time series information can be used to test causal relationships, which leads to a bigger sample, increases the number of degrees of freedom, and reduces the collinearity among explanatory variables, thus improving the efficiency of Granger causality tests (Hurlin and Venet, 2001; Hurlin and Dumitrescu, 2012). It also increases flexibility in the modeling of cross-sectional units' behavior compared to time series (Hood et al., 2008).

The methodology used to test the causal link between aid and government expenditures in this chapter is discussed in the next subsection.

Granger causality in the panel setting

To test for Granger causality in a heterogeneous panel setting, I follow the methodology first proposed by Hurlin and Venet (2001) and applied, among others, by Hood et al. (2008), Erdil and Yetkiner (2009), and Arvin and Barillas (2002). The notation stays close to these papers. The following fixed effects model is estimated:

$$y_{i,t} = \sum_{k=1}^{p} \alpha_k y_{i,t-k} + \sum_{k=k_0}^{p} \beta_k^i x_{i,t-k} + \gamma_i + u_{i,t} \tag{4.3}$$

where i denotes the recipient, k the lag and t the year, all autoregressive coefficients α_k and regression coefficients β_k^i are assumed to be constant over time for all k, and the parameters α_k are equal for all countries,[82] whereas β_k^i can be country specific. γ_i is the fixed country effect, $u_{i,t}$ denotes an error term. I test both instantaneous causality ($k_0 = 0$) and causality ($k_0 = 1$). As Hurlin and Venet (2001) point out, the only difference between instantaneous causality and causality procedures is the formulation of the null and alternative hypothesis.

[82] These assumptions are needed to retain sufficient degrees of freedom.

The first hypothesis to be tested is the Homogeneous Non Causality hypothesis (HNC) that all lagged values of x are zero.

$$H_0 : \beta_k^i = \beta_k^j = 0 \qquad \forall i \in [1, N], \forall k \in [k_0, p], i \neq j \tag{4.4}$$

$$H_1 : \beta_k^i \neq 0 \qquad \exists(i, k) \tag{4.5}$$

The following Wald statistic is calculated:[83]

$$F_{HNC} = \frac{(SSR_{r1} - SSR_u)/(N(p+a))}{SSR_u/[NT - N(1+(p+a)) - (p+a)]} \tag{4.6}$$

where $a = 1$ is used for instantaneous causality (additional restriction) and $a = 0$ for a simple causality, SSR_u denotes the sum of squared residuals of the unrestricted model described in the equation (4.3), and SSR_{r1} is the sum of squared residuals of the model under the null hypothesis. N is the number of countries, and p the number of lags. Rejection of the null hypothesis means that there is at least one country for which x is Granger-causing y.

If HNC hypothesis is rejected, I test whether causality is homogeneous for all recipients. The null hypothesis of the Homogeneous Causality hypothesis (HC) states that all β_k^i are equal for all countries. Formally:

$$H_0 : \beta_k^i = \beta_k^j \qquad \forall i, j \in [1, N], \forall k \in [k_0, p] \tag{4.7}$$

$$H_1 : \beta_k^i \neq \beta_k^j \qquad \exists(i, j, k) \tag{4.8}$$

Where SSR_{r2} denotes the sum of squared residuals under the null hypothesis. The Wald statistic is given by:

$$F_{HC} = \frac{(SSR_{r2} - SSR_u)/[(p+a)(N-1)]}{SSR_u/[NT - N(1+(p+a)) - (p+a)]} \tag{4.9}$$

When both homogeneous causality hypotheses (HNC and HC) are rejected, the process is heterogeneous. There may be a causal relation from x to y for a group of countries and no relation for the remaining countries.

To estimate the Heterogeneous Non Causality hypothesis (HeNC), a set of dummy variables for each country is created. These binary variables are

[83] Since the panel used in this study is unbalanced, T is the average number of periods in the sample.

multiplied by a vector of independent variables (either aid or government expenditures) to create a set of variables that assign the value of the independent variable for a particular country and zero for all other countries.[84] That gives a set of 53 variables for each lag used, one per lag for each country and leads to 53 tests. As Hurlin and Venet (2001) point out, in this case the cross sectional information is only used to improve the specification of the model and the power of tests. The null and alternative hypotheses under HeNC are:

$$H_0 : \beta_k^i = 0 \qquad \forall i \in [1, N], \forall k \in [k_0, p] \tag{4.10}$$

$$H_1 : \beta_k^i \neq 0 \qquad \forall i \in [1, N], \forall k \in [k_0, p] \tag{4.11}$$

The following Wald statistic is used:

$$F_{HeNC} = \frac{(SSR_{r3} - SSR_u)/(p+a)}{SSR_u/[NT - N(1+2(p+a)) + (p+a)]} \tag{4.12}$$

where SSR_{r3} is the sum of squared residuals under the null hypothesis. Using this specification, it can be detected for which countries there are causal links between aid and government expenditures.

A well known problem in studies that use a lagged dependent variable is dynamic panel bias, as was pointed out by Nickell (1981). However, the bias decreases with T (Judson and Owen, 1999), and for long panels a fixed effects estimator should work well (Roodman, 2009a). Therefore, the fixed effects model is used.

4.3 Data

The initial sample consists of 91 countries classified by the OECD as Least Developed Countries (LDCs) and Lower and Middle Income Countries (LMICs).[85] Data cover the period 1980–2009, and are (almost) complete starting from the 1990's. The use of lagged variables increases the minimum length of the

[84] Hood et al. (2008) describe this procedure in more detail.

[85] The OECD classification from 2009 is used. The term Least Developed Countries (LDCs) is used to describe any country with per capita GNI below $936 in 2007 (therefore it also includes Other Low Income Countries in this study). Lower Middle Income Countries (LMICs) have per capita GNI between 936 and 3705 in 2007. A more detailed description of the whole dataset, details on the data definitions, and a discussion of the quality and the precise sources can be found in Marć (2012).

time series. Hurlin and Dumitrescu (2012) use the formula $T > 5 + 2p$, p being the number of lags, to determine the number of time periods needed for estimation.[86] For 3 lags the minimum number of periods equals 12, and 14 when instantaneous causality is tested. Therefore, all countries with fewer than 14 observations for government expenditures or aid are dropped, which decreases the number of analyzed countries from 91 to 53,[87] and the number of country-year observations to 1145.

As noted before, this study focuses on the link of on-budget aid and government expenditures. The share of total government spending in GDP is taken from the International Monetary Fund's (IMF) World Economic Outlook (WEO). Data on aid disbursement are taken from the DAC2 OECD Table. Since precise data on on-budget and off-budget aid are not available, following Van de Sijpe (2013b) technical cooperation (TC) is used as a proxy for off-budget aid. Technical cooperation includes payments to consultants, advisers, and similar personnel in recipient countries and grants to citizens of recipient countries for education and training both at home and abroad, and its vast majority is not channeled through the budget. The remainder, i.e. total aid minus technical cooperation, is a proxy for on-budget aid. Admittedly, it still contains off-budget elements.[88]

Table 24 presents descriptive statistics. On average, on-budget aid accounts for 6.50% of recipient countries' GDP, 9.34% for the Least Developed Countries and 2.06% for the Lower and Middle Income Countries. When the amounts are weighted by the size of economy, the average drops to 0.65%, which reflects the fact that populous countries, like China and India, receive relatively small amounts of aid per capita. Also the discrepancy between LDCs and LMICs increases, the former group receives more than 10 times more aid expressed as a share of GDP. The share of government expenditures and the number of observations per country are stable across the groups.

[86] Hurlin and Dumitrescu (2012) point out that the second order moments of the Wald statistics exist only if this condition is satisfied.

[87] The list of the countries can be found in the Appendix.

[88] Foster (2005) studies 14 donors and finds that on average out of any 1$ of aid around 50 dollar cents are not recorded in the budget. For the health sector in Uganda and Tanzania, Gottret and Schieber (2006) estimate that around 50% of aid bypasses the budget.

Table 24: Share of government expenditures and on-budget aid in GDP (%)

	Unweighted		Weighted			
	Aid	Gov exp	Aid	Gov exp	Countries	Years
Sample	6.50	26.37	0.65	21.04	53	22.83
LDCs	9.34	26.01	4.61	20.17	32	23.12
LMICs	2.06	26.92	0.25	21.13	21	22.38

4.4 Results

This section presents results from the Granger causality tests. In the first part, tests for stationarity and the choice of lag length are discussed. Then, results for homogeneous and heterogeneous panel causality are presented and discussed.

Non-stationarity and the choice of lag length

Granger causality assumes weak stationarity in order for the results to be valid; the first and second moments should be constant. All variables used in the study are expressed in percentages, as a share of GDP, hence stationarity is likely. Both Im-Pesaran-Shin and Fisher tests reject the null hypothesis that all panels contain a unit root. Therefore, I assume that the variables are weakly stationary.

The choice of the lag length may affect the results of the Granger causality test. Odaki (1986) tests model selection criteria suggested by Akaike (1973), Schwarz (1978), and Hannan and Quinn (1979), pointing out desirable asymptotic properties. However, these properties may not hold in small samples and the conclusions reached with these procedures are conditional on the criteria and information used (Urbain, 1989). Thornton and Batten (1985) show that the arbitrary lag-length specification can lead to misleading results. They suggest, that "*the safest approach is to perform an extensive search of the lag space*" (p. 177). Arvin and Barillas (2002) list the consequences of the wrong choice of the lag lengths. If too many lags are used, it is likely that the power of the test is reduced. Conversely, if too few lags are used, the results will be biased and residuals will be serially correlated. To determine the optimal length of the lags both AIC and BIC criteria are used. The following fixed effects model for the pooled dataset is estimated for both aid and government expenditures:

$$y_{i,t} = \sum_{k=1}^{p} \alpha_k y_{i,t-k} + \gamma_i + \varepsilon_{i,t} \tag{4.13}$$

The results of both criteria suggest to use the model with the maximum number of lags available. Due to the length of time series (discussed in Section 4.3), which may lead to a low number of degrees of freedom, I limit the maximum number of lags analyzed to three.

Granger causality tests

After fixing the lag length at three, the procedures described in Section 4.2 are followed. Equation (4.3) is estimated in order to test the HNC and HC hypotheses (also for instantaneous causality). Table 25 presents the values of the Wald statistics for both hypotheses tested. The null of the Homogeneous Non Causality is rejected for both links and all groups at 1% level and the same applies to the Homogeneous Causality hypothesis. The rejections of HNC indicate that there is a causal relation from aid to government expenditures and from government expenditures to aid, and since the HC is also rejected, these causal links are not homogeneous.

In the next step, the heterogeneous causality hypothesis is tested for each country in the dataset, allowing for different slopes (β_k^i). The results are presented in Table 26 for both causality and instantaneous causality. For the causality from on-budget aid to government expenditures, at a 5% significance level six Least Developed Countries contribute to the existence of the causal link: Equatorial Guinea, Eritrea, Burundi, Sudan, Solomon Islands, and Maldives. Additionally, at 10% level Uzbekistan and Comoros are also in that group. Among the Lower and Middle Income Countries, only for Mongolia and Iran aid is Granger-causing government expenditures. Only for one country from LMICs, the Republic of Congo, government expenditures are Granger-causing aid, whereas it is the case for seven countries from the group of LDCs (and for one more at 10% level). All in all, for three countries the causality link is bidirectional: Equatorial Guinea, Burundi, and Eritrea.

Table 25: Homogeneous Non Causality Hypothesis (HNC) and Homogeneous Causality Hypothesis (HC) tests

Whole sample					
Hypothesis	Causality	Granger causality		Inst. Granger causality	
tested:		F	P-val	F	P-val
HNC	Aid GC GovExp	1.741	0.000	2.035	0.000
	GovExp GC Aid	1.729	0.000	3.657	0.000
HC	Aid GC GovExp	1.644	0.000	1.968	0.000
	GovExp GC Aid	1.728	0.000	3.692	0.000
Lower and Middle Income Countries (LMICs)					
Hypothesis	Causality	Granger causality		Inst. Granger causality	
tested:		F	P-val	F	P-val
HNC	Aid GC GovExp	1.393	0.003	1.500	0.000
	GovExp GC Aid	2.008	0.000	1.745	0.000
HC	Aid GC GovExp	1.323	0.009	1.454	0.000
	GovExp GC Aid	1.848	0.000	1.631	0.000
Least Developed Countries (LDCs)					
Hypothesis	Causality	Granger causality		Inst. Granger causality	
tested:		F	P-val	F	P-val
HNC	Aid GC GovExp	1.895	0.000	2.308	0.000
	GovExp GC Aid	1.762	0.000	4.025	0.000
HC	Aid GC GovExp	1.759	0.000	2.210	0.000
	GovExp GC Aid	1.711	0.000	4.014	0.000

There are 21 Lower and Middle Income Countries in the sample and 32 Least Developed Countries. Out of the 13 countries for which any causal relation is at the 5% level, only Mongolia, Iran, and the Republic of Congo are from the group of LMICs (14% of the group), the remaining 10 are Least Developed Countries (and account for 31% of that group). This is in line with expectations described in the introduction. Since aid accounts for only a small share of GDP (and consequently of government expenditures) among Lower and Middle Income Countries, the response of the level of government expenditures to the changes in aid is smaller (or not detectable). Also, LMICs have better access to credit markets and foreign aid may crowd out borrowing, hence an inconclusive answer is more likely than in more credit-constrained LDCs.

As a robustness check, I include five controls in equation (4.3): annual GDP growth (in %), lagged annual inflation rate, lagged agricultural value added as a share of GDP, lagged openness of the economy (sum of imports and exports of goods and services as a share of GDP), and lagged infant mortality rate per 1000 births. These regressors have been used in other studies to explain the size of government expenditures or inflow of aid.[89] At a 10% level, the causality link stays bidirectional for Equatorial Guinea and Eritrea. Aid is still found to Granger-cause government expenditures for Mongolia, Sudan, Uzbekistan, Solomon Islands and Iran, and additionally for Rwanda, while for Burundi and Maldives inconclusive results are found. For all other countries, the results remain inconclusive. For the causality going from government expenditures to aid, compared to the main results from Table 26 the conclusions change for two countries: Granger causality is found for Mongolia, while an inconclusive result for the Republic of Congo.

In summary, for the majority of countries, aid neither Granger-causes government expenditures, nor the other way around. That means that past values of aid (government expenditures) are not helpful in predicting the current level of government expenditures (aid). It further suggests two things. Firstly, the government reacts to on-budget aid by changing the way it uses its own resources, which means that aid is fungible. When aid flows increase, the government reduces its own revenues or net borrowing to keep the level of total government expenditures constant and therefore aid is substituting government expenditures. Secondly, aid does not respond to changes in government expenditures since past levels of government expenditures are not predictive for the current values of aid. This suggests that the disbursement schedule may be predetermined and that recipients do not have discretionary power to affect the timing of disbursement. That may also mean that donors do not react to the budgetary situation in the recipient country.

[89] Chapter two discussed the choice of these controls. Agricultural value added controls for spending in agriculture and serves as a proxy for the level of development. Infant mortality is a proxy for expenditures in social security and health sectors. The growth rate is used a proxy for shocks to expenditures: if recipients can decide on the timing of aid disbursement, the government can respond to shocks to expenditures by changing the amount of foreign aid. Openness of economy and inflation rate have been found to explain variation in the size of government expenditures in the literature.

Table 26: Granger causality and instantaneous Granger causality

LDCs		LMICs	
Country	Direction	Country	Direction
Equatorial Guinea	Bidirectional (also inst.)	Mongolia	Aid causing GovExp (also inst.)
Eritrea	Bidirectional (also inst.)	Iran	Aid causing GovExp (only inst.) *
Burundi	Bidirectional (also inst.)	Moldova	Inconclusive
Sudan	Aid causing GovExp (also inst.)	Morocco	Inconclusive
Uzbekistan	Inconclusive *,**	Algeria	Inconclusive
Solomon Islands	Aid causing GovExp (only inst.)*	Swaziland	Inconclusive
Maldives	Aid causing GovExp (only inst.)*	Rep. of Congo	GovExp causing Aid (only inst.)†
Lesotho	Inconclusive	China	Inconclusive
Chad	Inconclusive	Ecuador	Inconclusive
Comoros	Inconclusive **	El Salvador	Inconclusive
Djibouti	Inconclusive	Thailand	Inconclusive
Bhutan	Inconclusive	Jordan	Inconclusive
Ghana	Inconclusive	Colombia	Inconclusive
Papua New Guinea	Inconclusive	Sri Lanka	Inconclusive
Dem. Rep. of Congo	GovExp causing Aid (also inst.)	Namibia	Inconclusive
Mozambique	GovExp causing Aid (also inst.)	Syria	Inconclusive
Benin	Inconclusive	Bolivia	Inconclusive
Rwanda	GovExp causing Aid (also inst.)	Paraguay	Inconclusive
Guinea	Inconclusive	Tunisia	Inconclusive
Central African Rep.	Inconclusive	Philippines	Inconclusive
Cambodia	Inconclusive	India	Inconclusive
Burkina Faso	Inconclusive		
Tanzania	Inconclusive		
Yemen	Inconclusive		
Togo	Inconclusive		
Ethiopia	Inconclusive		
Vanuatu	Inconclusive † †		
Madagascar	GovExp causing Aid (only inst.)		
Pakistan	Inconclusive		
Kenya	Inconclusive		
Bangladesh	Inconclusive		
Niger	Inconclusive		

Countries are sorted according to the p-value of the test for Granger causality from aid to government expenditures. 5% level is assumed for inference.
* at 10% level, aid is Granger-causing government expenditures.
** at 10% level, aid is instantaneously Granger-causing government expenditures.
† at 10% level, also government expenditures are Granger-causing aid.
† † at 10% level, also government expenditures are instantaneously Granger-causing aid.

As noted before, recent studies on aid effectiveness show that aid has a positive impact on development outcomes, see for example: Arndt et al. (2010, 2011); Alvi and Senbeta (2012); Juselius et al. (2014); Mekasha and Tarp (2013). At the same time, at the aggregate level on-budget aid is partly fungible and around 40–50% substitutes government revenues in the long run (see chapter two) and, as this study shows, for the majority of countries on-budget aid is not Granger-causing government expenditures. These results suggest that it is quite plausible that the main impact of aid is not through increased government spending. Other channels may be also important. On-budget aid may have an impact on the structure and quality of the entire government activities, pushing for expenditures that are important for development. It may also happen that thanks to on-budget aid the tax burden is lower, and hence returns to private investment and resources available for private consumption are higher. What is more, it is also possible that non-fungible off-budget aid is driving the positive impact on growth and poverty reduction.

4.5 Conclusions

This study investigates the causal link between on-budget aid and government expenditures. A Granger causality test that allows for cross-sectional heterogeneity of recipients and heterogeneity in coefficients' slopes is applied to data for 53 developing countries for the period 1980–2009. For most countries, I find no evidence of Granger causality. For only six out of 32 Least Developed Countries on-budget aid Granger-causes government expenditures while for seven countries government expenditures Granger-cause aid. The two groups overlap: for three countries causality is bidirectional. For the Lower and Middle Income Countries, aid Granger-causes government expenditures in only two out of 21 countries, and in a single country the causality is in the opposite direction while no bidirectional causality was found. These results are reasonable as it is easier to make aid fungible in richer countries as it accounts for a small part of government expenditures. However, it is also quite likely that due to the relatively small amounts of aid causality may not be detectable for LMICs.

These findings suggest that (i) contrary to conventional wisdom on-budget aid does not increase expenditures (fungibility); most recipient governments adjust other sources of government revenues e.g. taxation or borrowing, (ii) donors do not react to the changes in government expenditures, presumably because there is little scope for changing the timing of disbursement. My findings indicate that the positive impact of aid on development outcomes presented in recent studies does not reflect the financing role of on-budget aid (the quantity effect), but changes in the use of government resources and reduced domestic resource mobilization (the quality effect), or the effectiveness of off-budget aid.

4.6 Appendix

Least Developed Countries: Bangladesh, Benin, Bhutan, Burkina Faso, Burundi, Cambodia, the Central African Republic, Chad, Comoros, the Democratic Republic of Congo, Djibouti, Equatorial Guinea, Eritrea, Ethiopia, Ghana, Guinea, Kenya, Lesotho, Madagascar, Maldives, Mozambique, Niger, Pakistan, Papua New Guinea, Rwanda, the Solomon Islands, Sudan, Tanzania, Togo, Uzbekistan, Vanuatu, Yemen.

Lower and Middle Income Countries: Algeria, Bolivia, China, Colombia, the Republic of Congo, Ecuador, El Salvador, India, Iran, Jordan, Moldova, Mongolia, Morocco, Namibia, Paraguay, the Philippines, Sri Lanka, Swaziland, Syria, Thailand, Tunisia.

CHAPTER 5

Conclusions

5.1 Summary

Foreign aid remains an important source of government revenues in developing countries. Therefore, the behavior of a recipient country's government may be crucial in understanding aid effectiveness. This PhD thesis focuses on the impact of development assistance on government expenditures and contributes to the fungibility and fiscal response literatures.

Chapter two investigates the impact of development assistance on government expenditures among 118 aid recipients in the period 1980–2012. First, I re-examine the fungibility of total aid at the aggregate level in the short- and long-term. Aid is found to be partly fungible: in the long-run, government expenditures increase by between 40% and 50% of aid, and the adjustment process is gradual as aid is highly fungible in the short run. There is no evidence that the response of government expenditures depends on whether aid is increasing or decreasing, which suggests that the governments are able to replace missing revenues to a large extent when aid is decreasing. It follows from the government budget constraint that when aid is fungible at the aggregate level, government adjusts domestic revenues or net borrowing, or that aid bypasses the budget. The results of aid's impact on government revenues indicate that own domestic revenues decrease in response to aid. In this chapter, I distinguish between off- and on-budget components of aid.

Technical cooperation is used as a proxy for off-budget aid. Since off-budget aid is not recorded in the budget, it is non-fungible if government expenditures do not change in response to aid. While on-budget aid is partly fungible, as expected, off-budget aid is found to be non-fungible. Moreover, the differences between bilateral and multilateral aid suggest that their impact on government expenditures may be different. I test this hypothesis and find that multilateral aid is less fungible than bilateral aid.

In chapter three potential endogeneity of aid is addressed. Aid may be endogenous in regressions explaining government expenditures due to at least three factors. Firstly, the recipient country's government may be allowed to determine the timing of aid disbursement and therefore the level of aid may depend on domestic revenues. Secondly, donors may target countries that fail to provide merit goods. Thirdly, some countries, possibly those with good policies, may be allowed to treat part of the aid as fungible. I use two instrumental methods from the fungibility literature. First, the difference and system GMM estimators are applied to the same dataset as was used in chapter two. I discuss advantages and disadvantages of these estimators and then apply these methods closely following the recommendation of Roodman (2009a). Aid is found to have a statistically insignificant impact on government expenditures in most specifications. However, this is not a conclusive result as I find that the instrument set is weak. Moreover, I show that unless these estimators are applied with great care, they can easily generate misleading results. In this chapter, I also replicate the results of Chatterjee et al. (2012) and show that their difference GMM estimate, and especially their system GMM, is not robust to changes in the instrument set. Second, an instrument set based on geographical and cultural proximity between the donor and the recipient is used. In the models that use dynamic specification, the aid coefficient is close to zero and insignificant. For the models without lagged government expenditures aid is statistically significant and the government expenditures increase by between 48% and 64% of aid. However, this instrument set is weak according to various specification tests, which suggests that the results may be biased.

Chapter four investigates the causal links between on-budget aid and government expenditures. I use Granger causality methods for panel data to investigate whether on-budget aid is useful in predicting government expenditures, and whether government expenditures are useful in predicting on-budget aid. Firstly, I investigate whether Granger causality does not exist in the whole sample using the Homogenous Non Causality (HNC) hypothesis. The HNC is rejected, which means that there exist at least one country where aid Granger-causes government expenditures and at least one country where government expenditures Granger-cause aid. In the next step, the Homogenous Causality (HC) hypothesis is tested: under the null hypothesis the causality pattern is homogenous among all countries. Not surprisingly, this hypothesis is also rejected, which means that the causality patterns differ among countries. Therefore, the Heterogenous Non Causality (HeNC) hypothesis is tested for each country. I find that for most countries neither on-budget aid Granger-causes government expenditures nor government expenditures Granger-cause on-budget aid. These results suggest that aid is fungible, and that donors do not respond to the budgetary situation in the recipient country. Furthermore, these results indicate that improvements in the use of government resources due to donor involvement or lower taxes and borrowing are likely channels of the impact of aid on development indicators, rather than the financing role of on-budget aid (the quantity effect).

In summary, in this PhD thesis I have used various econometric techniques to investigate the impact of aid on government expenditures using extensive recent dataset. Aid is found to be partly fungible and each additional euro of aid increases government expenditures by around 40 to 50 cents. Off-budget aid is found to be non-fungible, while on-budget aid partly fungible. Moreover, the fungibility of bilateral and multilateral aid differs. I believe that a potential bias coming from endogeneity of aid in this regression is small. However, I also show that two instrumental methods used in the fungibility literature are unlikely to provide reliable, unbiased results. Moreover, as expected in a large sample of countries, the impact of aid on government expenditures is heterogenous. Overall, these results suggest that part of aid is financing other projects than intended by the donor and that the amounts spent on earmarking are at least partly wasted. As noted in the introduction, fungibility is not necessarily bad for development. Recipient countries may allocate aid to more needed investments and aid may decrease taxes, which are often distortionary. Therefore, partial fungibility of aid means that at the macro level, aid's impact is a sum of increased spending, increased private consumption and savings (due to tax decreases), and potential improvements in the structure and quality of government expenditures due to the donor involvement.

5.2 Further research

As in many other applied fields of economics, there is an enormous discussion about instruments in aid effectiveness studies. However, there seems to be insufficient discussion about the data quality. The Development Assistance Committee (DAC) of the OECD is the main source of foreign aid data. Despite their big effort to harmonize donor practices and standardize reporting methodology, the dataset of foreign aid across countries is still not without problems. Aggregate aid consists of many different types of aid: grants and loans, cash transfers, goods etc. Therefore, the impact of aggregate aid, both in fungibility and aid-growth studies, reflects the influence of many different factors. This would indicate that disaggregated data should be used to estimate the impact of aid. However, while some dimensions of disaggregation are available (e.g. loans and grants, or bilateral and multilateral aid), many important aspects are missing. For example, as discussed in chapter two, there is no information on whether aid is channeled through the budget, or whether it bypasses the budget and directly reaches ultimate beneficiaries. Moreover, information on the sectoral allocation of aid is incomplete for many years in the Credit Reporting System of the OECD (see e.g. Van de Sijpe 2013b). Also, non-DAC members like China and Saudi Arabia are becoming increasingly important aid donors and the information on their activities is incomplete (see e.g. Manning 2006 or Dreher et al. 2010). Data on private donations are also largely incomplete, and for some donor countries, especially the US, private donations may account for a significant proportion of total development assistance. Furthermore, the data on government finances are of dubious quality (see e.g. Morrissey 2012). For example, the IMF WEO database contains only very limited information about tax revenues. Data on net borrowing and lending are not collected, but are calculated as a difference between estimated government expenditures and estimated government revenues (as discussed in chapter two). These data problems not only hinder investigation of various channels that could increase our understanding of aid effectiveness, but also increase the probability of reaching invalid conclusions because of the noise in the data that are available. Nevertheless, the quality of the aid data seems to be improving, for example the coverage of CRS data from the OECD has substantially increased, and new initiatives like AidData research and innovation lab (http://aiddata.org/) give hope that more comprehensive and reliable data will be available in the future. As I show in chapter three, the validity of earlier results from the literature should be tested with the most recent data and new econometric techniques. The case study literature describing budgetary processes in developing countries is mostly ignored in econometric studies of fungibility. The impact of aid on government expenditures depends on the information about the amount, composition, and timing of aid disbursement the recipient country's government possess. For example, the fungibility of aid may depend on the uncertainty with regard to the exact amount of aid disbursement in the coming year. It is likely that the models used in aid effectiveness studies miss potential channels of aid impact. Qualitative research could contribute to better understating the determination of government expenditures and the role the donors play, and hence it would be possible to improve existing theoretical and empirical models.

Bibliography

Akaike, H. (1973). Information Theory and an Extension of the Maximum Likelihood Principle. In Petrov, B. and F.Csaki, editors, *Second International symposium of information theory*, pages 267–281. Akademiai Kiado, Budapest.

Alesina, A. and Dollar, D. (2000). Who Gives Foreign Aid to Whom and Why? *Journal of Economic Growth*, 5(1):33–63.

Alesina, A. and Wacziarg, R. (1998). Openness, Country Size and Government. *Journal of Public Economics*, 69(3):305–321.

Alvi, E. and Senbeta, A. (2012). Does Foreign Aid Reduce Poverty? *Journal of International Development*, 24(8):955–976.

Andersen, T. G. and Sørensen, B. E. (1996). GMM Estimation of a Stochastic Volatility Model: a Monte Carlo Study. *Journal of Business & Economic Statistics*, 14(3):328–352.

Annen, K. and Kosempel, S. (2009). Foreign Aid, Donor Fragmentation, and Economic Growth. *The BE Journal of Macroeconomics*, 9(1):1–30.

Arellano, M. and Bond, S. (1991). Some Tests of Specification for Panel Data: Monte Carlo Evidence and an Application to Employment Equations. *Review of Economic Studies*, 58(1):277–297.

Arellano, M. and Bover, O. (1995). Another Look at the Instrumental Variable Estimation of Error-Components Models. *Journal of Econometrics*, 68(1):29–51.

Arndt, C., Jones, S., and Tarp, F. (2010). Aid, Growth, and Development: Have We Come Full Circle? *Journal of Globalization and Development*, 1(2):1–27.

Arndt, C., Jones, S., and Tarp, F. (2011). Aid Effectiveness: Opening the Black Box. Working Paper Seriers, World Institute for Development Economic Research (UNU-WIDER).

Arndt, C., Jones, S., and Tarp, F. (2014). Assessing Foreign Aid‘s Long Run Contribution to Growth and Development. *World Development*.

Arvin, B. M. and Barillas, F. (2002). Foreign Aid, Poverty Reduction, and Democracy. *Applied Economics*, 34(17):2151–2156.

Bachmeier, L. J. and Griffin, J. M. (2003). New Evidence on Asymmetric Gasoline Price Responses. *The Review of Economics and Statistics*, 85(3):772–776.

Bai, J. and Ng, S. (2010). Instrumental Variable Estimation in a Data Rich Environment. *Econometric Theory*, 26(06):1577–1606.

Bastiaens, J. (2007). *International Assistance and State-university Relations in Indonesia (1978–1998): Seeking Autonomy Within a Bureaucratic Polity*. State University of New York at Albany.

Baum, C. F., Schaffer, M. E., and Stillman, S. (2007). Enhanced Routines for Instrumental Variables/GMM Estimation and Testing. *Stata Journal*, 7(4):465–506.

Bazzi, S. and Clemens, M. A. (2013). Blunt Instruments: Avoiding Common Pitfalls in Identifying the Causes of Economic Growth. *American Economic Journal: Macroeconomics*, 5(2):152–186.

Berthélemy, J.-C. (2006). Bilateral Donors' Interest vs. Recipients' Development Motives in Aid Allocation: Do All Donors Behave the Same? *Review of Development Economics*, 10(2):179–194.

Bertrand, M., Duflo, E., and Mullainathan, S. (2004). How Much Should We Trust Differences-in-Differences Estimates? *The Quarterly Journal of Economics*, 119(1):249–275.

Blundell, R. and Bond, S. (1998). Initial Conditions and Moment Restrictions in Dynamic Panel Data Models. *Journal of Econometrics*, 87(1):115–143.

Bond, S. R., Hoeffler, A., and Temple, J. (2001). GMM Estimation of Empirical Growth Models. Technical report, Nuffield College (University of Oxford).

Boone, P. (1996). Politics and the Effectiveness of Foreign Aid. *European Economic Review*, 40(2):289–329.

Bourguignon, F. and Sundberg, M. (2007). Aid Effectiveness - Opening the Black Box. *American Economic Review Papers And Proceedings*, 97(2):316–321.

Bowsher, C. G. (2002). On Testing Overidentifying Restrictions in Dynamic Panel Data Models. *Economics Letters*, 77(2):211–220.

Bun, M. J. and Kiviet, J. F. (2006). The Effects of Dynamic Feedbacks on LS and MM Estimator Accuracy in Panel Data Models. *Journal of Econometrics*, 132(2):409–444.

Bun, M. J. and Windmeijer, F. (2010). The Weak Instrument Problem of the System GMM Estimator in Dynamic Panel Data Models. *The Econometrics Journal*, 13(1):95–126.

Burnside, C. and Dollar, D. (2000). Aid, Policies, and Growth. *American Economic Review*, 90(4):847–868.

Cameron, A. C. and Trivedi, P. K. (2005). *Microeconometrics: Methods and Applications*. Cambridge University Press.

Cashel-Cordo, P. and Craig, S. G. (1990). The Public Sector Impact of International Resource Transfers. *Journal of Development Economics*, 32(1):17–42.

Cassen, R. and associates (1994). *Does Aid Work?* Oxford: Clarendon Press.

Chatterjee, S., Giuliano, P., and Kaya, I. (2007). Where Has All the Money Gone? Foreign Aid and the Quest for Growth. IZA Discussion Papers 2858, Institute for the Study of Labor (IZA).

Chatterjee, S., Giuliano, P., and Kaya, I. (2012). Where Has All the Money Gone? Foreign Aid and the Composition of Government Spending. *The B.E. Journal of Macroeconomics*, 12(1):1–34.

Clemens, M. A., Radelet, S., Bhavnani, R. R., and Bazzi, S. (2012). Counting Chickens When They Hatch: Timing and the Effects of Aid on Growth. *The Economic Journal*, 122(561):590–617.

Clements, B. J., Gupta, S., Pivovarsky, A., and Tiongson, E. (2004). Foreign Aid and Revenue Response: Does the Composition of Aid Matter? In Clements, B. J., Gupta, S., and Inchauste, G., editors, *Helping Countries Develop: The Role of Fiscal Policy*. International Monetary Fund, Washington DC.

Clist, P. and Morrissey, O. (2011). Aid and Tax Revenue: Signs of a Positive Effect Since the 1980s. *Journal of International Development*, 23(2):165–180.

Collier, P., Guillaumont, P., Guillaumont, S., and Gunning, J. W. (1997). Redesigning Conditionality. *World Development*, 25(9):1399–1407.

Collier, P. and Hoeffler, A. (2007). Unintended Consequences: Does Aid Promote Arms Races? *Oxford Bulletin of Economics and Statistics*, 69(1):1–27.

Crivelli, E., Gupta, M. S., Muthoora, M. P. S., and Benedek, M. D. (2012). Foreign Aid and Revenue: Still a Crowding Out Effect? Technical Report 12/186, International Monetary Fund.

Dalgaard, C.-J., Hansen, H., and Tarp, F. (2004). On The Empirics of Foreign Aid and Growth. *Economic Journal*, 114(496):F191–F216.

Deaton, A. (2010). Instruments, Randomization, and Learning about Development. *Journal of Economic Literature*, 48(2):424–455.

Devarajan, S., Rajkumar, A. S., and Swaroop, V. (2007). What Does Aid to Africa Finance? In Lahiri, S., editor, *Theory and Practice of Foreign Aid*, pages 333–355. Emerald Group Publishing Limited.

Devarajan, S. and Swaroop, V. (2000). The Implications of Foreign Aid Fungibility for Development Assistance. In Gilbert, C. L. and Vines, D., editors, *The World Bank: Structure and Policies*, pages 196–209. Cambridge University Press.

Dollar, D. and Levin, V. (2006). The Increasing Selectivity of Foreign Aid, 1984-2003. *World Development*, 34(12):2034–2046.

Dreher, A., Nunnenkamp, P., and Thiele, R. (2010). Are 'New' Donors Different? Comparing the Allocation of Bilateral Aid between Non-DAC and DAC Donor Countries. KOF Working papers 10–255, KOF Swiss Economic Institute, ETH Zurich.

Erdil, E. and Yetkiner, I. H. (2009). The Granger-causality Between Health Care Expenditure and Output: a Panel Data Approach. *Applied Economics*, 41(4):511–518.

Feyzioglu, T., Swaroop, V., and Zhu, M. (1998). A Panel Data Analysis of the Fungibility of Foreign Aid. *World Bank Economic Review*, 12(1):29–58.

Fleck, R. K. and Kilby, C. (2010). Changing Aid Regimes? US Foreign Aid from the Cold War to the War on Terror. *Journal of Development Economics*, 91(2):185–197.

Foster, M. (2005). MDG Oriented Sector and Poverty Reduction Strategies. Health, Nutrition, and Population Discussion Paper, The World Bank.

Galiani, S., Knack, S., Xu, L. C., and Zou, B. (2014). The Effect of Aid on Growth: Evidence from a Quasi-Experiment. Technical Report 6865, World Bank.

Gamkhar, S. and Oates, W. (1996). Asymmetries in the Response to Increases and Decreases in Intergovernmental Grants: Some Empirical Findings. *National Tax Journal*, 49(4):501–512.

Gang, I. N. and Khan, H. A. (1990). Foreign Aid, Taxes, and Public Investment. *Journal of Development Economics*, 34(1-2):355–369.

Gottret, P. and Schieber, G. (2006). *Health Financing Revisited: A Practitioner's Guide*. World Bank.

Granger, C. W. J. (1969). Investigating Causal Relations by Econometric Models and Cross-Spectral Methods. *Econometrica*, 37(3):424–438.

Granger, C. W. J. (1980). Testing for Causality : A Personal Viewpoint. *Journal of Economic Dynamics and Control*, 2(1):329–352.

Gujarati, D. (2007). *Basic Econometrics With Data Cd*. McGraw-Hill Education (India) Pvt Limited.

Hannan, E. J. and Quinn, B. G. (1979). The Determination of the Order of an Autoregression. *Journal of the Royal Statistical Society. Series B (Methodological)*, 41(2):190–195.

Hansen, H. and Tarp, F. (2000). Aid Effectiveness Disputed. *Journal of International Development*, 12(3):375–398.

Hansen, H. and Tarp, F. (2001). Aid and Growth Regressions. *Journal of Development Economics*, 64(2):547–570.

Hauk, W. R. J. and Wacziarg, R. (2009). A Monte Carlo Study of Growth Regressions. *Journal of Economic Growth*, 14(2):103–147.

Hayakawa, K. (2009). First Difference or Forward Orthogonal Deviations-which Transformation Should be used in Dynamic Panel Data Models?: a Simulation Study. *Economics Bulletin*, 29(3):2008–2017.

Herzer, D., Nowak-Lehmann, F., Dreher, A., Klasen, S., and Martinez-Zarzoso, I. (2014). Comment on Lof, Mekasha, and Tarp (2014). *World Development*.

Hines, J. R. and Thaler, R. H. (1995). The Flypaper Effect. *Journal of Economic Perspectives*, 9(4):217–226.

Holmqvist, G. (2000). Fungibility Parameters—a Comment on Their Reliability and Policy Implications From an Aid Practitioner.

Holtz-Eakin, D., Newey, W., and Rosen, H. S. (1988). Estimating Vector Autoregressions with Panel Data. *Econometrica*, 56(6):1371–1395.

Hood, M. I., Kidd, Q., and Morris, I. L. (2008). Two Sides of the Same Coin? Employing Granger Causality Tests in a Time Series Cross-Section Framework. *Political Analysis*, 16(3):324–344.

Hurlin, C. and Dumitrescu, E. (2012). Testing for Granger Non-causality in Heterogeneous Panels. *Economic Modelling*, 29(4):1450–1460.

Hurlin, C. and Venet, B. (2001). Granger Causality Tests in Panel Data Models with Fixed Coefficients. mimeo, University Paris IX.

Judson, R. A. and Owen, A. L. (1999). Estimating Dynamic Panel Data Models: a Guide for Macroeconomists. *Economics Letters*, 65(1):9–15.

Juodis, A. (2013). A Note on Bias-Corrected Estimation in Dynamic Panel Data Models. *Economics Letters*, 118(3):435–438.

Juselius, K., Møller, N. F., and Tarp, F. (2014). The Long-Run Impact of Foreign Aid in 36 African Countries: Insights from Multivariate Time Series Analysis. *Oxford Bulletin of Economics and Statistics*, 76(2):153–184.

Kapetanios, G. and Marcellino, M. (2010). Factor-GMM Estimation with Large Sets of Possibly Weak Instruments. *Computational Statistics & Data Analysis*, 54(11):2655–2675.

Kilian, L. and Vigfusson, R. J. (2011). Are the Responses of the U.S. Economy Asymmetric in Energy Price Increases and Decreases? *Quantitative Economics*, 2(3):419–453.

Kirchgassner, G. and Wolters, J. (2007). *Introduction To Modern Time Series Analysis*. Springer.

Larraín, B. and Tavares, J. (2004). Does Foreign Direct Investment Decrease Corruption? *Cuadernos de Economía*, 41(123):199–215.

Lebovic, J. H. (1988). National Interests and US Foreign Aid: the Carter and Reagan Years. *Journal of Peace Research*, 25(2):115–135.

Leiderer, S. (2012). Fungibility and the Choice of Aid Modalities: The Red Herring Revisited. Working Paper Series, World Institute for Development Economic Research (UNU-WIDER).

Lof, M., Mekasha, T. J., and Tarp, F. (2014a). Aid and Income: Another Time-Series Perspective. *World Development*.

Lof, M., Mekasha, T. J., and Tarp, F. (2014b). Rejoinder to Herzer, Nowak-Lehmann, Dreher, Klasen, and Martinez-Zarzoso (2014). *World Development*.

Lutkepohl, H. (2009). *New Introduction To Multiple Time Series Analysis*. Springer.

Maddala, G. (2001). *Introduction to Econometrics*. Wiley.

Maizels, A. and Nissanke, M. K. (1984). Motivations for Aid to Developing Countries. *World Development*, 12(9):879–900.

Manning, R. (2006). Will 'Emerging Donors' Change the Face of International Co-operation? *Development Policy Review*, 24(4):371–385.

Marć, Ł. (2012). New Evidence on Fungibility at the Aggregate Level. *Tinbergen Institute Discussion Papers*, 12-083/2.

McGillivray, M. and Morrissey, O. (2000). Aid Fungibility in Assessing Aid: Red Herring or True Concern? *Journal of International Development*, 12(3):413–428.

McGillivray, M. and Morrissey, O. (2001a). Aid Illusion and Public Sector Behaviour. *The Journal of Development Studies*, 37(6):118–136.

McGillivray, M. and Morrissey, O. (2001b). Fiscal Effects of Aid. Working Paper Seriers Discussion Paper No. 2001/61, World Institute for Development Economic Research (UNU-WIDER).

McGillivray, M. and Morrissey, O. (2004). Fiscal Effects of Aid. In Addison, T. and Roe, A., editors, *Fiscal Policy for Development*. UNU-WIDER/Palgrave, Basingstoke.

McGillivray, M. and Ouattara, B. (2005). Aid, Debt Burden and Government Fiscal Behaviour in Cote d'Ivoire. *Journal of African Economies*, 14(2):247–269.

Mehrhoff, J. (2009). A Solution to the Problem of Too Many Instruments in Dynamic Panel Data GMM. Discussion Paper Series 1: Economic Studies 2009,31, Deutsche Bundesbank Research Centre.

Mekasha, T. J. and Tarp, F. (2013). Aid and Growth What Meta-Analysis Reveals. *The Journal of Development Studies*, 49(4):564–583.

Melitz, J. and Toubal, F. (2014). Native Language, Spoken Language, Translation and Trade. *Journal of International Economics*, 93(2):351–363.

Morrissey, O. (2004). Conditionality and Aid Effectiveness Re-evaluated. *The World Economy*, 27(2):153–171.

Morrissey, O. (2012). Aid and Government Fiscal Behaviour: What Does the Evidence Say? Working Paper Series, World Institute for Development Economic Research (UNU-WIDER).

Morrissey, O. (2014). Aid and Government Fiscal Behavior: Assessing Recent Evidence. *World Development*.

Nickell, S. J. (1981). Biases in Dynamic Models with Fixed Effects. *Econometrica*, 49(6):1417–1426.

Nowak-Lehmann, F., Dreher, A., Herzer, D., Klasen, S., and Martínez-Zarzoso, I. (2012). Does Foreign Aid Really Raise Per Capita Income? A Time Series Perspective. *Canadian Journal of Economics*, 45(1):288–313.

Odaki, M. (1986). Tests of Granger Causality by the Selection of the Orders of a Bivariate Autoregressive Model. *Economics Letters*, 22(2,3):223–227.

OECD (2001). Untying Aid to the Least Developed Countries. Policy brief, OECD.

OECD (2003). Harmonising Donor Practices for Effective Aid Delivery. Technical report, OECD.

OECD (2011). Update of Multilateral Agencies', Non-DAC Donors' and Private Founfactions' Statistical Reporting to the DAC in 2010. Technical Report DCD/DAC/STAT (2011) 21, OECD.

Osei, R., Morrissey, O., and Lloyd, T. (2005). The Fiscal Effects of Aid in Ghana. *Journal of International Development*, 17(8):1037–1053.

Pack, H. and Pack, J. R. (1990). Is Foreign Aid Fungible? The Case of Indonesia. *Economic Journal*, 100(399):188–194.

Pack, H. and Pack, J. R. (1993). Foreign Aid and the Question of Fungibility. *The Review of Economics and Statistics*, 75(2):258–265.

Pettersson, J. (2007a). Child Mortality: Is Aid Fungibility in Pro-Poor Expenditure Sectors Decisive? *Review of World Economics (Weltwirtschaftliches Archiv)*, 143(4):673–693.

Pettersson, J. (2007b). Foreign Sectoral aid Fungibility, Growth and Poverty Reduction. *Journal of International Development*, 19(8):1074–1098.

Radchenko, S. (2005). Oil Price Volatility and the Asymmetric Response of Gasoline Prices to Oil Price Increases and Decreases. *Energy Economics*, 27(5):708–730.

Rajan, R. G. and Subramanian, A. (2008). Aid and Growth: What Does the Cross-Country Evidence Really Show? *The Review of Economics and Statistics*, 90(4):643–665.

Ram, R. (2003). Roles of Bilateral and Multilateral Aid in Economic Growth of Developing Countries. *Kyklos*, 56(1):95–110.

Ram, R. (2004). Recipient Country's 'Policies' and the Effect of Foreign Aid on Economic Growth in Developing Countries: Additional Evidence. *Journal of International Development*, 16(2):201–211.

Ram, R. (2009). Openness, Country Size, and Government Size: Additional Evidence from a Large Cross-country Panel. *Journal of Public Economics*, 93(1-2):213–218.

Remmer, K. L. (2004). Does Foreign Aid Promote the Expansion of Government? *American Journal of Political Science*, 48(1):77–92.

Rodrik, D. (1998). Why do More Open Economies Have Bigger Governments? *The Journal of Political Economy*, 106(5):997–1032.

Roodman, D. (2007). The Anarchy of Numbers: Aid, Development, and Cross-Country Empirics. *World Bank Economic Review*, 21(2):255–277.

Roodman, D. (2008). Through the Looking-Glass, and What OLS Found There: On Growth, Foreign Aid, and Reverse Causality. Working Papers 137, Center for Global Development.

Roodman, D. (2009a). How to Do xtabond2: An Introduction to Difference and System GMM in Stata. *Stata Journal*, 9(1):86–136.

Roodman, D. (2009b). A Note on the Theme of Too Many Instruments. *Oxford Bulletin of Economics and Statistics*, 71(1):135–158.

Schwarz, G. (1978). Estimating the Dimension of a Model. *The Annals of Statistics*, 6(2):461–464.

Staiger, D. and Stock, H. (1997). Instrumental Variables Regression with Weak Instruments. *Econometrica*, 65:557–586.

Stine, W. F. (1994). Is Local Government Revenue Response to Federal Aid Symmetrical? Evidence from Pennsylvania County Governments in an Era of Retrenchment. *National Tax Journal*, 47(4):799–816.

Stock, J. H. and Yogo, M. (2005). Testing for Weak Instruments in Linear IV Regression. In *Identification and Inference for Econometric Models: Essays in Honor of Thomas Rothenberg*, volume 1, pages 80–108. Cambridge University Press.

Swaroop, V., Jha, S., and Sunil Rajkumar, A. (2000). Fiscal Effects of Foreign Aid in a Federal System of Governance: The Case of India. *Journal of Public Economics*, 77(3):307–330.

Tarp, F. (2006). Aid and Development. *Swedish Economic Policy Review*, 13:9–61.

Tavares, J. (2003). Does Foreign Aid Corrupt? *Economics Letters*, 79(1):99–106.

Thornton, D. L. and Batten, D. S. (1985). Lag-Length Selection and Tests of Granger Causality between Money and Income. *Journal of Money, Credit and Banking*, 17(2):164–178.

Todaro, M. and Smith, S. (2003). *Economic Development*. Prentice Hall.

Urbain, J. P. (1989). Model Selection Criteria and Granger Causality Tests : An Empirical Note. *Economics Letters*, 29(4):317–320.

Van de Sijpe, N. (2013a). The Fungibility of Health Aid Reconsidered. *Journal of Development Studies*, 49(12):1746–1754.

Van de Sijpe, N. (2013b). Is Foreign Aid Fungible? Evidence from the Education and Health Sectors. *The World Bank Economic Review*, 27(2):320–356.

Van de Walle, D. and Cratty, D. (2005). Do Donors Get What They Paid For? Micro Evidence on the Fungibility of Development Project Aid. Policy Research Working Paper Series 3542, The World Bank.

Van de Walle, D. and Mu, R. (2007). Fungibility and the Flypaper Effect of Project Aid: Micro-evidence for Vietnam. *Journal of Development Economics*, 84(2):667–685.

Wagstaff, A. (2011). Fungibility and the Impact of Development Assistance: Evidence from Vietnam's Health Sector. *Journal of Development Economics*, 94(1):62–73.

Windmeijer, F. (2005). A Finite Sample Correction for the Variance of Linear Efficient Two-step GMM Estimators. *Journal of Econometrics*, 126(1):25–51.

Younas, J. (2008). Motivation for Bilateral Aid Allocation: Altruism or Trade Benefits. *European Journal of Political Economy*, 24(3):661–674.

Samenvatting (Dutch Summary)

Ontwikkelingshulp is nog steeds een belangrijke bron van overheidsinkomsten in ontwikkelingslanden. Om de effectiviteit van deze hulp te begrijpen is het essentieel om naar het gedrag van de overheid van een ontvangend land te kijken. In dit proefschrift pas ik verschillende econometrische technieken toe om de invloed van ontwikkelingshulp op overheidsinkomsten en -uitgaven te onderzoeken. Hiermee bepaal ik in hoeverre hulp fungibel is, wat wil zeggen dat de overheidsuitgaven met minder dan het volledige bedrag aan hulp stijgen. Dit kan komen doordat de overheid van het ontvangende land de binnenlandse inkomsten of leningen aanpast, of als hulp buiten de begroting om gaat.

Hoofdstuk twee van het proefschrift onderzoekt het effect van ontwikkelingshulp op de overheidsuitgaven van 118 hulpontvangende landen in the periode 1980-2012. Ten eerste onderzoek ik nogmaals de fungibiliteit van het totaalbedrag aan hulp voor zowel de korte als de lange termijn. Op de lange termijn nemen de overheidsuitgaven met 40% tot 50% van het bedrag aan hulp toe. Dit is echter een geleidelijk proces: op de korte termijn is hulp zeer fungibel. Voor het gevonden effect op overheidsuitgaven maakt het niet uit of de hulp op- of afgebouwd wordt, wat suggereert dat overheden in staat zijn om afnemende hulpinkomsten grotendeels te vervangen door andere inkomsten. Uit de budgetrestrictie van de overheid volgt dat wanneer hulp op het hoogste niveau fungibel is, de overheid de binnenlandse inkomsten of het netto geleende bedrag aanpast, óf dat de hulp buiten de begroting om gaat. De resultaten wijzen erop dat de eigen binnenlandse inkomsten dalen als gevolg van de hulp. Ik maak in dit hoofdstuk onderscheid tussen hulp die via de begroting gaat, en hulp die buiten de begroting om gaat. Hierbij wordt technische samenwerking tussen landen gebruikt als proxy voor hulp buiten de begroting om. Aangezien hulp die buiten de begroting om gaat per definitie niet is opgenomen in de begroting, is deze hulp niet fungibel als blijkt dat de overheidsuitgaven gelijk blijven als reactie op de hulp. Zoals verwacht is hulp die via de begroting gaat gedeeltelijk fungibel, maar hulp buiten de begroting blijkt niet fungibel te zijn. Verder zou er door het verschil een karakter van bilaterale en multilaterale hulp ook een verschil kunnen zijn in het effect dat zij hebben op de overheidsuitgaven. Ik test deze hypothese en vind dat multilaterale hulp minder fungibel is dan bilaterale hulp.

In hoofdstuk drie bespreek ik de mogelijke endogeniteit van hulp. Er zijn drie redenen waarom ontwikkelingshulp mogelijk endogeen is in regressies die overheidsuitgaven verklaren. In de eerste plaats, bepaalt de overheid van het ontvangende land mogelijk de timing van de uitbetaling van de hulp, waardoor de verkregen hoeveelheid hulp afhankelijk kan worden van de overheidsinkomsten. Ten tweede kan het zijn dat donorlanden hulp bieden aan landen die niet in staat zijn om de bevolking te voorzien van *merit goods*. Ten derde kunnen ontvangende landen die kunnen aantonen goed beleid te voeren in sommige gevallen een deel van de hulp als fungibel mogen behandelen. Voor de endogene variabele hulp gebruik ik daarom twee instrumentele variabelen die gebruikelijk zijn in de literatuur over fungibiliteit. De eerste methode past de *difference* en *system GMM* schatters toe op de dataset die ook in hoofdstuk twee gebruikt wordt. Ik beschrijf de voor- en nadelen van de schatters en pas ze vervolgens toe, waarbij ik de aanbevelingen van Roodman (2009) volg. In de meeste specificaties blijkt het effect van hulp op overheidsuitgaven niet significant. Dit is echter nog geen definitieve conclusie, omdat de gebruikte instrumenten zwak blijken te zijn. Ik laat zien dat als deze schatters niet met zorg worden toegepast, ze makkelijk tot misleidende resultaten kunnen leiden. In dit hoofdstuk repliceer ik ook de resultaten van Chatterjee (2012). Ik laat zien dat de door hen toegepaste *difference GMM* schatter, en met name de *system GMM* schatter, gevoelig zijn voor veranderingen in de verzameling van instrumenten. De tweede verzameling van instrumenten die ik toepas is gebaseerd op de geografische en culturele nabijheid tussen het donorland en het ontvangende land. In de modellen met een dynamische specificatie is de coëfficiënt van hulp ongeveer nul en niet significant. In een model met daarin voorgaande waardes van overheidsuitgaven is het effect van hulp statistisch significant en verhoogt het de overheidsuitgaven met 48% tot 64% van het bedrag aan hulp. Uit verschillende specificatie testen blijkt echter dat de instrumenten set zwak is, wat kan betekenen dat de resultaten onzuiver zijn.

Hoofdstuk vier onderzoekt de causale link tussen hulp die via de begroting gaat en overheidsuitgaven. Ik gebruik hiervoor de Granger-causaliteit methode voor paneldata om te kijken of hulp via de begroting de hoogte van overheidsuitgaven kan verklaren, en of overheidsuitgaven de hoeveelheid hulp via de begroting kan verklaren. Allereerst kijk ik of er Granger-causaliteit aanwezig is in de hele dataset, waarbij ik gebruik maak van de Homogene Geen Causaliteitshypothese (HNC). Deze hypothese wordt verworpen, wat betekent dat er ten minste één land is waar hulp een Granger-causaal verband heeft met overheidsuitgaven en dat er ten minste één land is waar overheidsuitgaven een Granger-causaal verband hebben met hulp. De volgende stap is om de Homogene Causaliteitshypothese (HC) te testen. De nulhypothese is hierbij dat het patroon van de causaliteit gelijk is voor alle landen. Deze hypothese wordt zoals verwacht verworpen, wat betekent dat het patroon van de causaliteit verschilt tussen landen. Daarom wordt vervolgens voor elk land apart de Heterogene Geen Causaliteitshypothese (HeNC) toegepast. Ik concludeer dat voor de meeste landen hulp via de begroting geen Granger-oorzakelijk verband heeft met overheidsuitgaven en dat ook overheidsuitgaven geen Granger-oorzakelijk verband hebben met hulp via de begroting. Deze resultaten wijzen erop dat hulp fungibel is en dat donorlanden niet reageren op de budgettaire situatie in het ontvangende land. Verder tonen de resultaten aan dat verbeteringen in het gebruik van overheidsmiddelen als gevolg van bemoeienis van het donorland of lagere belastingen en lagere leenkosten mogelijke mechanismes zijn waarop hulp effect heeft op indicatoren voor de ontwikkeling van een land. Het financieringsaspect van de hulp speelt hierbij veel minder een rol.

Samengevat heb ik in dit proefschrift verschillende econometrische technieken toegepast om de invloed van hulp op overheidsuitgaven te onderzoeken. Ik heb hierbij gebruik gemaakt van een recente en zeer uitgebreide dataset. Het blijkt dat hulp gedeeltelijk fungibel is; elke extra euro hulp verhoogt de overheidsuitgaven in het ontvangende land met ongeveer de helft. Hulp buiten de begroting om is niet fungibel, terwijl hulp via de begroting gedeeltelijk fungibel is. Verder is er een verschil in fungibiliteit van bilaterale en multilaterale hulp. Ik ben er van overtuigd dat de potentiële onzuiverheid van de schatters als gevolg van endogeniteit van hulp klein is. Ik laat daarnaast zien dat twee veelgebruikte instrumenten in de fungibiliteitsliteratuur niet altijd tot zuivere en betrouwbare resultaten leiden. Ten slotte laat ik zien dat, zoals te verwachten is in een sample met veel landen, de invloed van hulp op de overheidsuitgaven verschilt tussen landen. Alles tezamen laten deze resultaten zien dat een deel van het geld van ontwikkelingshulp besteed wordt aan andere projecten dan de donor bedoeld had, en dat de middelen inzetten om geld te oormerken in ieder geval gedeeltelijk weggegooid geld zijn. Zoals hierboven al vermeld, is fungibiliteit niet per definitie slecht voor de ontwikkeling van een land. Landen die hulp ontvangen kunnen deze hulp gebruiken om belangrijke investeringen te doen of om (eventueel verstorende) belastingen te verlagen. Als er dan op macroniveau sprake is van gedeeltelijke fungibiliteit van hulp, is het effect van hulp een combinatie van hogere overheidsuitgaven, hogere particuliere consumptie en toegenomen spaargelden (als gevolg van belastingverlagingen) en van potentiële verbeteringen in de structuur en kwaliteit van de overheidsuitgaven door inmenging van het donorland.

The Tinbergen Institute is the Institute for Economic Research, which was founded in 1987 by the Faculties of Economics and Econometrics of the Erasmus University Rotterdam, University of Amsterdam and VU University Amsterdam. The Institute is named after the late Professor Jan Tinbergen, Dutch Nobel Prize laureate in economics in 1969. The Tinbergen Institute is located in Amsterdam and Rotterdam. The following books recently appeared in the Tinbergen Institute Research Series:

558 S. PEER, *The economics of trip scheduling, travel time variability and traffic information*
559 V. SPINU, *Nonadditive Beliefs: From Measurement to Extensions*
560 S.P. KASTORYANO, *Essays in Applied Dynamic Microeconometrics*
561 M. VAN DUIJN, *Location, choice, cultural heritage and house prices*
562 T. SALIMANS, *Essays in Likelihood-Based Computational Econometrics*
563 P. SUN, *Tail Risk of Equidity Returns*
564 C.G.J. KARSTEN, *The Law and Finance of M&A Contracts*
565 C. OZGEN, *Impacts of Immigration and Cultural Diversity on Innovation and Economic Growth*
566 R.S. SCHOLTE, *The interplay between early-life conditions, major events and health later in life*
567 B.N.KRAMER, *Why don't they take a card? Essays on the demand for micro health insurance*
568 M. KILIÇ, *Fundamental Insights in Power Futures Prices*
569 A.G.B. DE VRIES, *Venture Capital: Relations with the Economy and Intellectual Property*
570 E.M.F. VAN DEN BROEK, *Keeping up Appearances*
571 K.T. MOORE, *A Tale of Risk: Essays on Financial Extremes*
572 F.T. ZOUTMAN, *A Symphony of Redistributive Instruments*
573 M.J. GERRITSE, *Policy Competition and the Spatial Economy*
574 A. OPSCHOOR, *Understanding Financial Market Volatility*
575 R.R. VAN LOON, *Tourism and the Economic Valuation of Cultural Heritage*
576 I.L. LYUBIMOV, *Essays on Political Economy and Economic Development*
577 A.A.F. GERRITSEN, *Essays in Optimal Government Policy*
578 M.L. SCHOLTUS, *The Impact of High-Frequency Trading on Financial*

Markets

579 E. RAVIV, *Forecasting Financial and Macroeconomic Variables: Shrinkage, Dimension reduction, and Aggregation*

580 J. TICHEM, *Altruism, Conformism, and Incentives in the Workplace*

581 E.S. HENDRIKS, *Essays in Law and Economics*

582 X. SHEN,*Essays on Empirical Asset Pricing*

583 L.T. GATAREK, *Econometric Contributions to Financial Trading, Hedging and Risk Measurement*

584 X. LI, *Temporary Price Deviation, Limited Attention and Information Acquisition in the Stock Market*

585 Y. DAI, *Efficiency in Corporate Takeovers*

586 S.L. VAN DER STER, *Approximate feasibility in real-time scheduling: Speeding up in order to meet deadlines*

587 A. SELIM, *An Examination of Uncertainty from a Psychological and Economic Viewpoint*

588 B.Z. YUESHEN, *Frictions in Modern Financial Markets and the Implications for Market Quality*

589 D. VAN DOLDER, *Game Shows, Gambles, and Economic Behavior*

590 S.P. CEYHAN, *Essays on Bayesian Analysis of Time Varying Economic Patterns*

591 S. RENES, *Never the Single Measure*

592 D.L. IN 'T VELD, *Complex Systems in Financial Economics: Applications to Interbank and Stock Markets*

593 Y.YANG, *Laboratory Tests of Theories of Strategic Interaction*

594 M.P. WOJTOWICZ, *Pricing Credits Derivatives and Credit Securitization*

595 R.S. SAYAG, *Communication and Learning in Decision Making*

596 S.L. BLAUW, *Well-to-do or doing well? Empirical studies of wellbeing and development*

597 T.A. MAKAREWICZ, *Learning to Forecast: Genetic Algorithms and Experiments*

598 P. ROBALO, *Understanding Political Behavior: Essays in Experimental Political Economy*

599 R. ZOUTENBIER, *Work Motivation and Incentives in the Public Sector*

600 M.B.W. KOBUS, *Economic Studies on Public Facility use*

601 R.J.D. POTTER VAN LOON, *Modeling non-standard financial decision making*
602 G. MESTERS, *Essays on Nonlinear Panel Time Series Models*
603 S. GUBINS, *Information Technologies and Travel*
604 D. KOPÁNYI, *Bounded Rationality and Learning in Market Competition*
605 N. MARTYNOVA, *Incentives and Regulation in Banking*
606 D.KARSTANJE, *Unraveling Dimensions: Commodity Futures Curves and Equity Liquidity*
607 T.C.A.P. GOSENS, *The Value of Recreational Areas in Urban Regions*