# Does bilateral aid work better through multilateral channels? Evidence from the GODAD database

Bernhard Reinsberg[1,2]

1 – School of Social and Political Sciences, University of Glasgow

2 – Centre for Business Research, University of Cambridge

Correspondence: 42 Bute Gardens, Glasgow, G12 8RT, bernhard.reinsberg@glasgow.ac.uk

**Abstract**: A key decision for donors is whether to delegate aid to a multilateral organization. While much literature has sought to understand the drivers of aid delegation, less is known about its effectiveness. While principal–agent theory expects delegation benefits that make aid more effective, recent studies have established the costs of narrow delegation, specifically a loss of performance of multilateral organizations where donors delegate strictly earmarked funding. To empirically examine the effectiveness of earmarked aid in comparison with bilateral aid, I turn to the novel GODAD database, which provides geolocated information on aid projects from 22 official donors from 1989 to 2024. I combine the GODAD database with the Earmarked Funding Dataset to identify bilateral projects involving multilateral implementers and the level of delegated autonomy in such projects. Using subnational analysis over 1,753 first-level administrative regions from 1990 to 2022, I deploy several inferential strategies to examine the effect of different types of project funding on subnational development outcomes—household wealth, child mortality, and governance quality—available from the Global Data Lab. The analysis suggests that bilateral projects and earmarked projects have heterogenous impacts across these development outcomes. There are also differences in the effectiveness of different types of earmarked aid. The results hold important implications for debates in aid effectiveness, particularly regarding the effectiveness of aid implementation modalities.

JEL codes: F35, F53, O19

**Acknowledgements**: I am grateful to participants of the GODAD workshop on 12-13 March 2025 in Göttingen for helpful comments. Funding from UK Research & Innovation (Future Leaders Fellowship – MR/V022148/1) is gratefully acknowledged.

## 1. Introduction

In an increasingly dense global development architecture, donors have many possibilities for how to deliver their aid. A key decision involves the delegation of aid implementation to a multilateral organization, such as the World Bank, the United Nations, and regional economic institutions. According to principal−agent theory, such acts of delegation entail significant benefits to states, given the professional expertise, global reach, and policy credibility of multilateral organizations. At the same time, delegation incurs costs to states, notably the loss of control (Abbott and Snidal 1998; Hawkins et al. 2006; Nielson and Tierney 2003; Rodrik 1995).

An open question is whether and how delegation influences aid effectiveness. Theoretical predictions from the principal−agent framework are ambiguous. On the one hand, delegation may increase aid effectiveness because multilaterals offer economies of scale and their mandates are tilted toward advancing pro-poor development (Annen and Knack 2015; Milner and Tingley 2013; Reinsberg et al. 2017). Empirical studies tend to confirm that multilateral aid is generally more effective than bilateral aid (Dreher et al. 2024; Headey 2008; Minoiu and Reddy 2010). On the other hand, recent studies on earmarked funding in international organizations find that delegated aid might be ineffective if donors attach too many strings, for example with earmarked contributions (Heinzel, Cormier, and Reinsberg 2023; Heinzel, Reinsberg, and Zaccaria 2024; Reinsberg and Siauwijaya 2024). Political economy literature also finds that donors may undermine aid effectiveness when meddling with the ordinary governance process of multilateral organizations in the pursuit of geopolitical objectives, even though politically motivated interventions into multilateral aid policies are the exception rather than the rule (Dreher et al. 2013; Kaya et al. 2021; Kilby 2013; Stone 2011).

Clear-cut predictions have further been muddied by the increasing variety of aid delegation contracts (Baumann et al. 2019; Martens et al. 2003; Reinsberg et al. 2023). On the one hand, donors may afford multilaterals with wide discretion even when using earmarked funding instruments. Pooled funds supported by multiple donors—such as the UN SDG Fund—provide a case in point. On the other hand, donors may restrict the use of their funds narrowly to specific project-type interventions. In fact, the bulk of earmarked funding to the UN system comes as narrowly earmarked aid activities (Baumann et al. 2019). A prominent argument is that more restrictive funds limit the autonomy of the multilateral agency, which may hinder development effectiveness, especially in contexts where flexibility is key (Honig 2019).

Rather than assuming that any specific aid modality is unconditionally more effective than another, I argue that their relative effectiveness is context-dependent. My theoretical framework proposes two key factors that determine which approach works best: the substantive fit of the aid intervention with the procedures of the multilateral bureaucracy, and the extent to which states face a credibility problem in aid delivery. Based on these factors, I argue that different aid modalities are better suited for different development goals. Specifically, bilateral aid—given directly from one country to another—is more effective at promoting economic growth but less effective at improving governance quality. Delegated aid—given through a multilateral organization—is more effective at improving governance quality, especially when donors tie their hands and afford multilaterals with greater autonomy to make credible policy commitments, using softly earmarked funding mechanisms. Delegated aid that limits the discretion of multilateral organizations can still be effective in situations where a quick response is needed, such as in the case of communicable diseases.

For an empirical test of these expectations, I turn to the novel GODAD database, which provides geolocated information on aid projects from 22 official donors from 1989 to 2024 (Bomprezzi et al. 2024). I combine this database with the Earmarked Funding Dataset to identify bilateral projects involving multilateral implementers and the type of delegation contract in such projects (Reinsberg et al. 2023). I consider three metrics of effectiveness, tapping different dimensions of subnational development impact. The first is the *International Wealth Index*, an asset-based measure of household wealth that captures the level of socio-economic development in a given region. The second is *under-five child mortality*, a widely used measure of development in the area of health. The third is the *Subnational Corruption Index* (Crombach and Smits 2024), which captures the quality of subnational governance through the absence of petty corruption and grand corruption. All three metrics are drawn from the Global Data Lab, combining multiple individual-level surveys to create comparable metrics of subnational development (Smits 2016).

Using panel data at the ADM-1 level—the political-administrative level where all measures are available—I seek to isolate the development impacts of foreign aid using variants of fixed-effects estimations, some including pre-weighted observations to eliminate the confounding influence of pre-treatment trends. I also deploy an instrumental-variable design with shift-share instruments, which are constructed through the interaction between the likelihood of aid of a given subnational unit *and* the total budget of aid of a given modality in all subnational locations in a given world region in a given year. The findings suggest that earmarked projects and bilateral projects have heterogenous impacts across different dimensions of subnational development. In particular, earmarked aid tends to fare better than bilateral aid when it comes to tackling child mortality and building state capacity, but fares worse than bilateral aid with respect to generating wealth. There are also differences in the effectiveness of different types of earmarked aid. Whereas strictly earmarked aid reduces child mortality, only softly earmarked aid enhances governance quality. These results demonstrate the complexities of how aid modalities affect whether aid is effective, suggesting that neither modality is unconditionally better. Across different modalities, however, the overall contribution of aid to development impacts appears to be rather small, and few results hold up across different inferential designs.

The results hold important implications for debates in aid effectiveness, particularly regarding the differential effectiveness of aid implementation modalities. With respect to the debate about the effectiveness of earmarked aid, my analysis contributes novel insights by comparing the effectiveness of earmarked aid relative to bilateral aid. In fact, most studies have assumed that the natural counterfactual to earmarked aid is multilateral aid (Heinzel et al. 2023; Heinzel et al. 2024a; Heinzel and Reinsberg 2024). This assumption may be implausible to the extent that earmarked aid comes out of bilateral aid budgets for many DAC donors (Eichenauer and Reinsberg 2017; Tortora and Steensen 2014; Weinlich et al. 2020). Unlike multilateral aid, which is often locked in for several years for a given funding cycle, earmarked aid can be programmed in a more flexible manner, much like bilateral aid.

My analysis is the first to systematically investigate how different types of earmarked aid affect development outcomes. While a growing body of research has explored how earmarked funding impacts the performance of international organizations and their alignment with recipient-country priorities (Heinzel et al. 2024b), there is a significant gap in research on its actual development impacts. The only existing study on this topic focuses narrowly on subnational economic growth and compares earmarked projects and core-funded projects at the World Bank (Heinzel and Reinsberg 2024). My research expands on this by examining the earmarked activities of all multilateral organizations, not just the World Bank, and their impact on three different development

outcomes, while addressing a critical knowledge gap about the differential effectiveness of strictly earmarked aid *versus* softly earmarked aid (Reinsberg and Taggart 2025). This distinction is especially important because global initiatives, such as the UN Funding Compact, recommend that donors shift toward softly earmarked funding, believing it to be more effective (Baumann and Weinlich 2020; UN-MPTFO 2022; UN 2019). As softly earmarked funds are expected to become more prevalent, understanding their development effectiveness compared to other funding types is crucial.

A large literature examines when and why aid is effective (Doucouliagos and Paldam 2009; Dreher et al. 2024; Dreher and Lohmann 2015). This literature has generally reached mixed conclusions, arguing that researchers need to consider heterogeneous effects. While scholars agree that aid motives matter (Dreher et al. 2015, 2024; Dreher and Kilby 2010), seemingly technocratic decisions such as whether and how to delegate aid implementation can have unintended consequences (Ouattara and Strobl 2008; Winters 2019; Wright and Winters 2010). My work thus closely aligns with recent work demonstrating the performance cost of donor earmarking inside multilateral agencies, while taking a novel approach by comparing earmarked aid to non-delegated bilateral aid as underlying counterfactual.

## 2. Aid effectiveness and delegation to multilaterals

Whether foreign aid is truly effective is a major concern (Doucouliagos and Paldam 2009; Dreher et al. 2024; Gisselquist and Tarp 2019). *Effectiveness* can be understood in two main ways. First, in the most common and narrow sense, aid is effective if it helps a country achieve its development goals, such as those outlined in the Sustainable Development Goals (SDGs). This includes reducing poverty, improving health and education, and building more capable government institutions. Second, in a broader sense, aid can also be considered effective if it advances the policy goals of the donor country. These goals are often seen as secondary to the core mission of development and can include promoting economic ties with the recipient country, discouraging migration, and gaining political influence (Dreher et al. 2024).

While cross-country data has been the traditional focus of research on aid effectiveness, scholars are now using other methods to analyze the impact of foreign aid. One prominent area of research uses project evaluation scores to determine what makes a project successful, considering factors like project design, institutional features, and country-level characteristics (Ashton et al. 2023; Bulman et al. 2017; Chauvet et al. 2010; Eilers et al. 2025). Although project evaluations can tell if a project met its specific goals, they often do not show if a project had actual development impact. To address this, researchers are now using geocoded aid data, which maps where aid is being deployed. By combining this data with socioeconomic indicators and political outcomes in the same geographic areas, they can more accurately measure the tangible effects of aid. For instance, a growing body of work uses this geospatial approach to examine whether aid leads to development (Bitzer and Gören 2024; Bluhm et al. 2025; Cruzatti et al. 2023; Dreher et al. 2021; Dreher and Lohmann 2015) or promotes peace and stability (Bitzer et al. 2025; Brazys et al. 2017; Dreher et al. 2025; Gehring et al. 2022; Rustad et al. 2019). These studies have significantly improved our understanding of the local effects of different aid donors. However, a common limitation is that they frequently overlook the different modalities or methods donors use to deliver aid, which can also influence the outcome.

Donors can choose from different ways to deliver aid, reflecting the diverse goals they want to achieve. The traditional distinction is between bilateral aid and multilateral aid. In *bilateral aid*, donors give aid directly to a recipient-country government. This approach gives the donor full control and allows for a rapid response to new opportunities. It is useful for both quick humanitarian responses and for creating political goodwill with recipient governments. In *multilateral aid*, donors pool their resources and delegate aid implementation to an international organization, like the World Bank or the United Nations. This type of aid is typically governed by long-term agreements, making it less flexible for new issues. Its main purpose is to support global public goods over a longer period (Dreher et al. 2021; Milner and Tingley 2013; Radelet 2006).

In recent decades, a third method for delivering aid, known as *earmarked aid*, has become increasingly popular (Eichenauer and Reinsberg 2017; Graham 2023; Reinsberg et al. 2024). This type of aid involves donors giving money to multilateral organizations, but with specific conditions on how it can be used—for example, for a particular country, sectors, or project (Bayram and Graham 2017; Eichenauer and Reinsberg 2017; OECD 2011). Unlike multilateral aid contributions, earmarked aid often bypasses the formal governing bodies of these organizations. Instead, it creates a direct line of accountability between the donor and the multilateral organization. This approach allows for more flexibility than traditional multilateral aid but is less flexible than direct bilateral aid, especially when donors choose to support 'multi-donor trust funds' in which multiple donors are bound by the same governing framework. Earmarked funding is often used by small groups of donors—minilateral coalitions—to support development initiatives that might not have broad support from all member states. In some cases, a single donor will partner with a multilateral organization to create a special fund, using so-called 'single-donor trust funds' (Reinsberg et al. 2017). This gives the donor significant influence over how the funds are used without needing to maintain their own large-scale aid bureaucracy.

A recurrent point of theoretical debate is whether the way aid is delivered affects how effective it is (Dreher et al. 2024). Some believe that multilateral aid is more effective than bilateral aid with respect to promoting development. The literature on this topic often highlights three main advantages of using multilateral organizations. First, multilateral organizations can pool resources from multiple donors, leading to larger-scale projects and potentially greater impact (Milner and Tingley 2013). This holds not only for core funding but also for large multi-donor trust funds (Barakat et al. 2012). Second, working with a single multilateral organization can be more efficient than coordinating with multiple bilateral donors. Donors often are poorly informed about the activities of other donors on the ground, leading to duplication of efforts and overstretched recipient-country bureaucracies which need to deal with hundreds of individual donor missions every year (Arnab Acharya et al. 2006; Easterly and Pfutze 2008; Knack and Rahman 2007). Multilateral organizations can play a key role in coordinating aid efforts and harmonizing donor policies, reducing duplication and improving overall effectiveness. Third, multilateral organizations often have specialized knowledge and experience in specific areas, such as health, education, or infrastructure development (Heinzel 2022; Ravallion 2016; Weaver 2008). They are widely perceived as more neutral and impartial than bilateral donors, potentially leading to greater trust and acceptance in recipient countries (Eckhard and Parizek 2022). Empirical studies on aid effectiveness find some support that multilateral aid is more effective (in the narrow sense defined earlier) than bilateral aid. Panel analysis at the country-level finds that multilateral aid has a positive effect on economic growth, whereas bilateral aid failed to stimulate growth during the Cold War when geo-strategic rationales prevailed (Headey 2008). This result suggests that multilaterals are

more effective because they disburse ‘developmental aid’, which unlike ‘non-developmental aid’ is more effectively targeted toward promoting economic development (Minoiu and Reddy 2010).

Others believe the merits of delegation to multilateral organizations are likely exaggerated—even if for entirely different reasons. Organizational theorists criticize tight donor controls as they can undo the gains from delegation. These scholars argue that multilaterals require some degree of independence from their donor states to fulfill their roles (Bauer and Ege 2016; Campbell 2018; Honig 2019; Lall 2017). How donors choose to fund multilateral organizations directly impacts the autonomy of those organizations. While core funding allows multilaterals to use the funds as they see fit, in line with their mandates, earmarked funding, restricts how they can use the money. The level of restriction varies. Strictly earmarked contributions can severely limit organizational discretion, even if donors cannot influence all decisions, especially operational choices like project procurement. Softly earmarked contributions are more flexible and preserve a higher degree of autonomy. Recent research emphasizes how donor practices like earmarked funding can rapidly eliminate the benefits of delegation, thereby undermining aid effectiveness. By forcing multilaterals to accept short-term projects that are often not well-aligned with their core mandates, donors may stretch the capacities of multilaterals and set incentives for their staff to prioritize fundraising over programmatic work (Heinzel et al. 2024b; Reinsberg 2023; Schmid et al. 2021). Earmarked projects, especially if tightly earmarked, often do not focus on the neediest (Reinsberg 2023). They also pose additional operational challenges across the entire project cycle: the voluntary nature of these funds makes them less predictable than core funding, preventing sustained interventions that can exert impact on the ground. During project implementation, the additional reporting requirements that donors impose on their earmarked contributions increase the transaction cost of aid delivery, thereby reducing the amount of funding available for generating impact on the ground (Heinzel et al. 2023).

From a public-choice perspective, multilateral aid is not more effective than bilateral aid. This view holds that multilateral institutions, like any bureaucracy, are primarily self-interested (Frey 1997; Niskanen 1971; Vaubel 2006). They prioritize their own budget growth and internal procedures over their stated goal of promoting development. Organizational sociologists have made similar observations about bureaucratic challenges. However, they view these problems as more structural—inherent to the organizational design and bureaucratic procedures—rather than being driven by the self-interested behavior of specific individuals. They argue that multilateral organizations are prone to dysfunction because they are entrapped in rules and procedures that undermine their ability to respond rapidly to emerging needs (Barnett and Finnemore 2004; Lipson 2010; Weaver 2008). Multilaterals are also seen as being far removed from both the ultimate recipients of aid and the taxpayers who fund it, which creates a significant accountability problem (Lang 2020; Moravcsik 2004; Vaubel 2006). From these perspectives, the benefits of delegating aid to a multilateral organization quickly evaporate. Bypassing multilateral organizations could even reduce bureaucratic hurdles and streamline the aid delivery process, potentially leading to faster and more efficient implementation of projects. Where bilateral aid goes directly to the beneficiary government, the donor can also more easily monitor how the funds are used and hold the recipient government accountable for results. Moreover, bilateral aid can help foster political ties between the donor government and the recipient government. This may promote bilateral cooperation beyond aid, which in turn my enhance the productivity of aid. Finally, bilateral aid can be adjusted quickly to changing political realities on the ground. While this flexibility may disrupt aid in some cases, it may help using aid in the most effective way in other cases.

The above debate unveils seemingly irreconcilable claims. Advocates of delegation, who believe that multilateral organizations are more effective, argue that multilaterals need more discretion to be able to fulfil their mandates. In contrast, those who criticize multilaterals and who favor more control in the hands of states perceive the autonomy afforded to multilaterals as a problem, given that it furthers agency slack and ineffective development interventions. Both sides of the debate likely miss the point. Rather than assuming that any specific aid modality is unconditionally more effective than another, I argue that their relative effectiveness is context-dependent. In other words, both views may be correct, but only under some circumstances. The relative effectiveness of different aid modalities is determined by two factors: the substantive fit of the aid intervention with the procedures of the multilateral bureaucracy, and the extent to which donors have a credibility problem in aid delivery. By analyzing these factors, my framework expects variation in the effectiveness of different aid modalities across different development objectives.

Bilateral aid, which flows directly from one country to another, is expected to be relatively effective in promoting economic development. This is because bilateral donors can respond to emerging economic opportunities in a relatively unconstrained manner. Furthermore, this modality is politically savvy for both donor governments and recipient governments. Donors can use bilateral aid to achieve specific foreign policy objectives, while recipient governments often see economic growth as a key to their own political survival (Bueno de Mesquita and Smith 2009; Licht 2010). However, this same political expediency is what makes bilateral aid less effective for promoting institutional reforms and better governance. Donors cannot credibly commit to insisting on genuine governance reforms when their foreign policy interests might be better served by continuing to provide aid, even in the face of corruption or institutional weakness (Cheeseman et al. 2024; Dunning 2004; Swedlund 2017) .

To overcome this credibility deficit, delegated aid, particularly in the form of earmarked funding, can be a more effective mechanism for supporting governance reforms. By delegating some decision-making power to a multilateral organization, donors constrain their own ability to backtrack on reform conditions (Rodrik 1995). The effectiveness of this approach, however, depends on the level of discretion afforded to the multilateral. When donors provide softly earmarked funding, multilaterals gain the autonomy to make credible policy commitments and adapt to local governance challenges. Empirical research in political economy supports the notion that where donors interfere with ordinary multilateral processes, aid effectiveness decreases (Dreher et al. 2013; Heinzel and Reinsberg 2024; Kilby 2015). This suggests that delegation can enable donors to realize effectiveness gains with respect to governance outcomes. This expectation is consistent with the historical record: some bilateral donors utilize ‘bypass aid’, channeling funds through non-governmental organizations or local agencies instead of central governments. Delegation can ensure that aid still reaches its intended beneficiaries and achieves specific development outcomes, even when the donor is unwilling or unable to enforce robust governance reforms on the state itself (Dietrich 2021; Eichenauer and Reinsberg 2017; Wright and Winters 2010).

Conversely, strictly earmarked aid can be especially effective in situations where multilateral organizations might otherwise be slow to respond. This type of funding is ideal for issues requiring a quick and targeted approach, such as combating communicable diseases. Multilaterals often have an incentive to avoid these kinds of interventions because they are highly quantifiable (like child vaccinations), which makes it easier to expose organizational inefficiency or a lack of results. By avoiding them, multilaterals can sidestep accountability and maintain their exclusive right to define what constitutes effective development. In these cases, multilaterals know that donors will quickly

cut off funding if results are not delivered rapidly, especially since this aid is not tied up in long-term commitments like core contributions. Research confirms that donors use earmarked contributions to reward good performance but will also reduce this aid when multilaterals fail to deliver results (Reinsberg et al. 2024), even if their internal processes are efficient (Reinsberg and Siauwijaya 2024). Contrary to general expectations in the literature, I therefore expect that strictly earmarked aid is more effective in tackling communicable diseases than softly earmarked aid. Historical evidence is consistent with this expectation: Following growing frustrations with how the World Health Organization addressed the burgeoning HIV/AIDS epidemic, donors established 'vertical funds' that uprooted the ways in which health aid tackling communicable diseases was delivered. Bypassing broad-based health systems, these global funds instead targeted specific diseases with greater speed and performance-based financing, channeling incremental amounts of strictly earmarked funds into existing multilaterals (Fidler 2016; Gatti 2024; Sridhar and Woods 2013).

In sum, I have argued that the relative effectiveness of different aid modalities is context-dependent. Differences exist not only between bilateral aid and earmarked aid, but also among different types of earmarked aid, depending on how restrictive it is. Strictly earmarked aid allows donors to retain more control, which increases their temptation to misuse control for advancing non-developmental objectives, or pursuing development activities that have limited relevance to ultimate beneficiaries. Limiting the discretion of multilaterals through tighter control may be useful in areas where multilateral organizations have failed to act or where they eschew accountability. In contrast, softly earmarked aid removes some control from donors and allows multilaterals to use discretion in allocating the funds as they see fit. This delegated discretion can have benefits for development, especially in the area of governance aid where bilateral donors would otherwise have difficulties to make credible commitments.

Importantly, I remain agnostic as to whether aid through different delivery modalities is effective. As this issue is already the subject of a large aid effectiveness literature, I focus here on the differential effectiveness of delivery modalities, specifically bilateral aid versus earmarked aid and different types of earmarked aid.

## 3. Data and methods

To examine whether and how multilateral delegation of aid activities influences aid effectiveness, I construct a subnational panel dataset at the ADM1-level. The data covers 1,753 first-level administrative regions in 160 countries from 1990 to 2022. ADM1-level analysis is the most common type of subnational analysis in aid effectiveness research (Asmus-Bluhm et al. 2024; Dreher et al. 2021; Dreher and Lohmann 2015; Gehring et al. 2022; Wellner et al. 2025). The attraction of this data structure is that subnational regions are nested within countries, which allows for inferential strategies that include country-level fixed effects. Moreover, administrative-level data has become more widely available, thereby helping to mitigate confounding factors. The main challenge with subnational administrative units is that their boundaries are not drawn randomly and potentially correlated with both aid flows and key outcomes, requiring additional inferential strategies.

### 3.1. Outcome variables

For the outcomes of interest, I rely on the Global Data Lab, which has the most extensive coverage of subnational development outcomes across several dimensions. To construct the most comprehensive subnational development outcomes database, the Global Data Lab combines various surveys, including the Demographic and Health Surveys (DHS), the UNICEF Multiple Indicator Cluster Surveys (MICS), IPUMS-International, Afrobarometer, Arab Barometer Surveys, LAPOP Surveys, and other data sources (Smits 2016). I employ three outcome measures from the Global Data Lab.

First, the International Wealth Index (IWI) considers housing characteristics, ownership of certain durable consumer goods, and access to vital services like electricity and water supply, as proxies for household wealth. It is an asset-based index for measuring the economic status of households in developing countries (Smits and Steendijk 2015). By providing a stable yardstick for comparing households across the developing world, the IWI addresses a major problem with traditional wealth indices, which are often specific to a single survey and not comparable across different countries and over time. I use the average IWI value for a given subnational region in a given year, which can range from zero (low-quality housing and absence of assets) to one-hundred (high-quality housing and full access to all assets).

Second, I measure under-five child mortality rates, specifically the number of children dying under five years of age per 1,000 live births in a given year. Child mortality is a widely used indicator capturing aid effectiveness with respect to basic health (Cruzatti et al. 2023; Daoud and Reinsberg 2019; Kotsadam et al. 2018; Martorano et al. 2020).

Third, to measure governance outcomes, I use the Subnational Corruption Index (SCI). Higher values of this index indicate more robust governance, characterized by lower incidences of petty corruption and grand corruption (Crombach and Smits 2024).

### 3.2. Key predictors

For aid activity data, I draw on the Geolocated Official Development Assistance Database (GODAD). The dataset includes 909,212 annual observations of 236,651 distinct development projects from 22 donors from 1989 to 2024 (Bomprezzi et al. 2024). Data availability of other key variables of interest implies that the effective number of observations is smaller. As my interest is in the differential impact of different implementation modalities, I only consider aid projects from the 19 OECD/DAC donor governments in the GODAD sample to make observations more comparable.[1] I thus discard projects from China, India, and the World Bank but use the information about their co-located aid projects to construct control variables.

To identify multilaterally delegated aid activities of OECD/DAC donors, I draw on the Earmarked Funding Dataset (Reinsberg et al. 2023). This dataset includes 342,812 aid activities from up to 50 donors with over 340 multilaterals from 1990 to 2020. In cases where CRS data did not report a unique identifier for multilateral implementers, the Earmarked Funding Dataset identifies earmarked activities using a combination of keyword-assisted searches in project descriptions and implementing channels. While earmarked activities can be targeted toward a specific location, they can also have no geographic remits or no geographical earmarking at all. In total, I identify 17,788 (geolocated) earmarked projects. Of these, 73% are with UN entities, 17% are with the World Bank,

[1] In addition, earmarked funding is common only among bilateral DAC donors, suggesting that an analysis of these three additional donors would not be feasible.

8% are with regional development banks, and the remainder is with other multilaterals—including public-private partnerships and vertical funds.

A key benefit of the Earmarked Funding Dataset is to capture the stringency of delegation of any earmarked activity. In general, earmarked activities can be restricted in their geographical scope, thematic focus, or institutional modality. Since this analysis focuses on projects with already-restricted geographical locations, I assess earmarking stringency based on the thematic focus of an activity. I combine the systematic coding on named trust funds in the Earmarked Funding Dataset and a keyword search for program-based assistance to identify softly earmarked activities. As discussed further in the appendix, activities are generally assumed to be strictly earmarked unless their descriptions make explicit reference to a (known) trust fund or a thematic program within the multilateral institution (Box A1 in the appendix).

A key issue pertains to the aggregation of aid projects to the ADM-1 level. Three measurement choices are available. First, I construct *binary indicators* for whether an administrative region had any projects, distinguishing between (undelegated) bilateral projects and earmarked projects. The dummy-variable specification is advantageous as it is a robust measure that can be readily interpreted. Second, a common choice is to count the *number of projects* of each type, and taking the natural logarithm to mitigate skewness. Counts appear to be more precise but can induce measurement error to the extent that reporting practices are heterogenous across donors. Specifically, some donors provide more detailed reporting of individual activities, whereas others bundle activities into larger projects. Yet, because I aggregate project counts across all DAC donors, I can mitigate the bias arising from heterogeneous reporting. Third, a final option for aggregation is to calculate the *total disbursement* from all projects of a given type. This measurement choice addresses concerns about bias due to different reporting granularity but hinges on accurate reporting of project volumes. Using project volumes assumes that each dollar spent exerts the same effect on outcomes across different project types. Nonetheless, I consider disbursements as meaningful complements to the binary and count measures of project support.

### 3.3. Estimation approaches

In the baseline estimations, I use two-way fixed effects, respectively controlling for arbitrary time-invariant characteristics of administrative regions and (common) annual shocks. To insulate against omitted-variable bias, I seek to control for the most plausible confounders at the subnational level. Specifically, I include binary indicators for whether a region has any Chinese development finance projects and any World Bank projects (Bomprezzi et al. 2024; Custer et al. 2021). In addition, I include national-level measures of development, notably the Human Development Index (HDI) and the State Capacity Index (SCI). Both national measures are available from the Global Data Lab (Smits 2016). Variable definitions and descriptive statistics are shown in the appendix (Table 1).

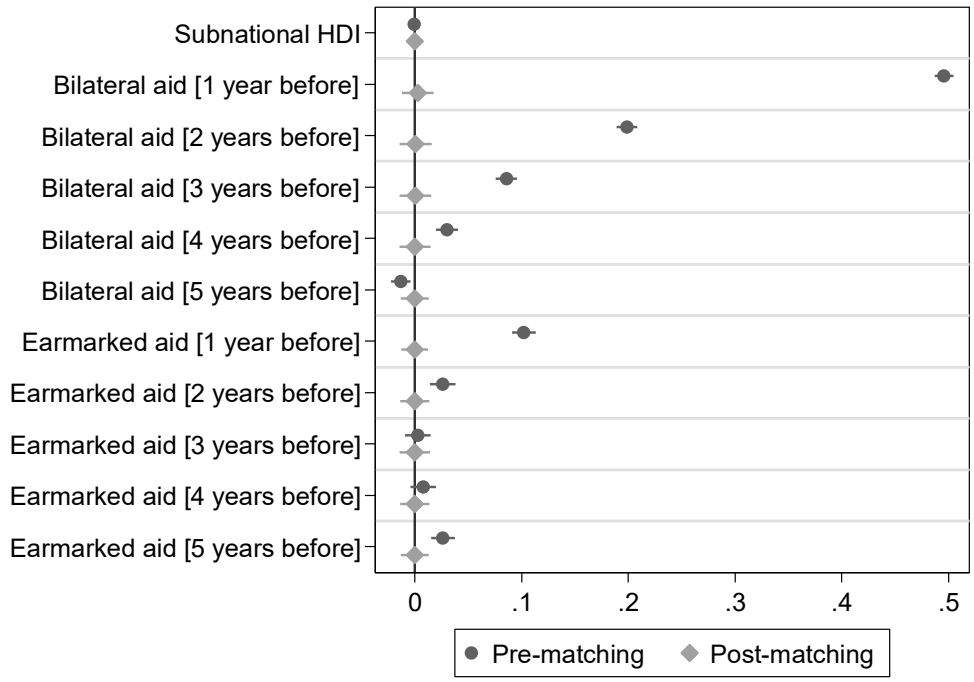
**Table 1**: Variable definitions and descriptive statistics

| Label | Description | Count | Mean | Sd | Min | Max |
|---|---|---|---|---|---|---|
| International Wealth Index (IWI) | The International Wealth Index (IWI) captures the extent to which households in a subnational area possess a bundle of goods and is expressed in percentage of the global maximum, available from the Global Data Lab (Smits 2016) | 27023 | 50.937 | 24.683 | 0.010 | 98.800 |
| Child mortality | Under-five child mortality refers to the number of children per 1,000 children that die before reaching the age of five years, available from the Global Data Lab (Smits 2016) | 24958 | 76.132 | 57.855 | 0.000 | 518.000 |
| Subnational Corruption Index (SCI) | Subnational Corruption Index (SCI) captures the extent to which subnational regions are free from petty corruption and grand corruption, based on household surveys on bribery and corruption, taken from the Global Data Lab (Crombach and Smits 2024) | 47488 | 56.778 | 12.558 | 12.700 | 92.700 |
| Bilateral aid | Binary indicator if a region has any bilateral project (from any DAC donor, excluding China, India, and World Bank) in a given year, based on GODAD Database (Bomprezzi et al. 2024) | 57849 | 0.415 | 0.493 | 0.000 | 1.000 |
| Earmarked aid | Binary indicator if a region has any project delegated to a named multilateral (from any DAC donor, excluding China and India) in a given year, based on Earmarked Funding Dataset (for earmarked activity) (Reinsberg et al. 2024) and GODAD Database (Bomprezzi et al. 2024) | 57849 | 0.093 | 0.290 | 0.000 | 1.000 |
| Softly earmarked aid | Binary indicator if a region has any softly earmarked aid project with a named multilateral. In a country-specific project, soft earmarking implies earmarking for a broad theme | 57849 | 0.038 | 0.192 | 0.000 | 1.000 |
| Strictly earmarked aid | Binary indicator if a region has any strictly earmarked aid project with a named multilateral. In a country-specific project, strict earmarking implies earmarking for a specific intervention, possibly also involving staff secondment | 57849 | 0.074 | 0.262 | 0.000 | 1.000 |
| Chinese projects | Binary indicator if a region has any Chinese projects in a given year, based on GODAD Database (Bomprezzi et al. 2024) | 57849 | 0.084 | 0.277 | 0.000 | 1.000 |
| World Bank projects | Binary indicator if a region has any World Bank (core) projects in a given year, based on GODAD Database (Bomprezzi et al. 2024) | 57849 | 0.438 | 0.496 | 0.000 | 1.000 |
| National HDI | National measure of the Human Development Index, from Global Data Lab (Smits and Permanyer 2019) | 53075 | 64.809 | 17.022 | 21.200 | 96.500 |
| National SCI | National measure of the State Capacity Index, from Global Data Lab (Crombach and Smits 2024) | 45441 | 56.887 | 12.135 | 20.200 | 92.200 |

To confront remaining inferential challenges, I use variants of high-dimensional fixed effects models and additional methods. First, I augment the two-way fixed effect model with country-period fixed effects.[2] This enables me to use control variables sparsely, ensuring limited data loss due to listwise deletion and mitigating omitted-variable bias. In line with common practice in the literature, the use of higher-dimensional fixed effects obviates the need for (many) substantive control variables. The high-dimensional fixed-effects specification identifies the relationship between the presence of different aid modalities and subnational outcomes relative to the country-period average. The key identifying assumption for a causal interpretation of the parameters of interest is that development outcomes in a given region would have evolved the same way regardless of whether or not it had any aid projects. This assumption is untestable. However, it is more likely to hold to the extent that non-project regions are similar to project regions.

To enhance the plausibility of the parallel-trends assumption, I adopt a pre-weighting approach to the data (Ahlfeldt et al. 2019). I employ entropy balancing and re-weight observations to generalize to a hypothetical population in which control regions and treatment regions have the same pre-treatment trends (Hainmueller 2012). This weighting approach seeks to ensure that administrative regions with bilateral projects and earmarked projects do not differ in their pre-treatment history and their initial development outcomes. To this end, I create indicators for whether a region had a given type of aid projects for the five years prior to a disbursing aid project. I then combine these indicator variables with the continuous subnational HDI to generate observation weights that should make treated observations and control observations alike. The subnational HDI is an omnibus measure of human development, which combines local estimates of education, health, and income into a single index (Smits and Permanyer 2019). Figure 1 confirms that the pre-weighting approach successfully achieves this.

**Figure 1**: Balance in covariates across project observations and non-project observations




[2] I tried to include country-year fixed effects but this would have prevented me to estimate IV/2SLS models due to too many variables (N>2,048). For the sake of consistency, I therefore used country-period effects allowing for country-specific heterogeneity across ten-year periods (1992-2001, 2002-2011, and 2012-2022).

*Notes*: The plot shows coefficient estimates from a linear-probability model predicting whether a region had any disbursing aid project in a given year, before matching (black circles) and after matching (gray diamonds). Coefficient estimates that are insignificant indicate that treated regions and control regions are perfectly balanced in the covariates after appropriate weights from the matching are being applied.

In additional tests, I draw on an inferential strategy that exploits variation in the timing of projects at the regional level (Isaksson and Kotsadam 2018). Specifically, I measure if a region had any projects of a given implementation modality that were planned but that did not yet disburse. Included jointly with the indicators for disbursing projects, these indicators remove any bias due to selection assuming that the omitted variables driving this selection do not change rapidly over time. This strategy thus promises to identify the causal effect of project disbursement when the time interval between project approval and project disbursement is reasonably short. In fact, the modal temporal distance between these two decisions in the data is just two years.

A final strategy for addressing inferential threats is an instrumental-variable design, to purge the choice of implementation modality of its endogenous variation. Specifically, I predict the presence of a project using the interaction between the average (time-invariant) probability of aid in the subnational region *and* the (time-varying) annual probability of aid across all the regions but within the same world region. This decomposition has been shown to yield plausibly exogenous instruments (Borusyak et al. 2022). Here, I assume that the region-specific probability of aid receipt reflects historic contingencies, such as colonial relations, which must be considered as endogenous. However, yearly trends in different types of aid funding reflect common shocks—such as newly emerging issues in global development and domestically driven fluctuation in donor budgets—that are unrelated to specific events in a given recipient-country subnational region. I therefore posit that the interaction of both variables is plausibly exogenous. The interaction instrument is strong ($F$>169), suggesting no issues with weak instruments.

To illustrate the intuition for the instrument, consider the case of aid to Africa in the past decade. When former German aid minister Gerd Müller took office in December 2013, he not only announced an increase in total aid but also sought stronger support for Africa, including through a 'Marshall Plan with Africa'—launched in 2017—which prioritized job creation, vocational training, and private sector engagement in African countries. This renewed ambition for Africa particularly benefited the earmarked aid budget, which increased from US$ 172 million to US$ 2 billion in 2013-2021. Conversely, total aid to Africa across all implementing channels increased from US$ 4.3 billion to US$ 7.3 billion in the same period (OECD 2022). The example shows the benefit of distinguishing modality-specific aid budgets to specific world regions for obtaining strong instruments.

## 4. Results

This section presents my results, examining the relationship between project funding and subnational development outcomes using various inferential approaches. I first present my main results, followed by robustness checks, and a final section addressing remaining threats to inference.

### 4.1. Main results

Table 2 shows the results from OLS regressions of three subnational development indicators on different types of project funding. For each outcome, I estimate three models. The first is a naïve two-way fixed effects model at the ADM-1 level. The second model augments this specification with country-period fixed effects, thereby accounting for medium-term changes in any confounders at the national level. The third model employs observation weights that seek to eliminate difference in pre-treatment trends. In all models, I control for the presence of World Bank projects and Chinese development finance. I also measure the concurrent values of the Human Development Index and the State Capacity Index at the national level. For inference purposes, I cluster standard errors at the administrative region.

As the estimates in table 2 indicate, the relationship between aid and development varies across aid modalities and development outcomes considered. In particular, while bilateral aid tends to increase household wealth, earmarked aid significantly reduces it, and the difference in coefficients is statistically significant ($p<0.01$). In substantive terms, based on the second model, the presence of a bilateral project is related to an increase in the International Wealth Index by 0.26 percentage points. An earmarked project is related to a decrease in the IWI by 0.26 percentage point—about 1.1 percent of its standard deviation (SD). These are small magnitudes, considering that an increase in the national HDI by 1 SD is related to an increase in the IWI by 9.53 percentage points (or 0.39 SDs).

With respect to infant mortality, the image is reversed: while bilateral aid appears to have no effect, earmarked aid tends to be negatively related to infant mortality—whenever it is statistically significant. However, coefficient magnitude and statistical significance rapidly decline as the model specification gets more demanding. This offers limited evidence to suggest that earmarked aid reduces infant mortality in subnational regions against the country trend, even if it is true that in regions with more earmarked aid, infant mortality decreases. Even under the most generous specification, coefficient magnitudes are relatively moderate: if a region has at least one earmarked aid project, its infant mortality is predicted to decrease by at most 3.88 children—about 7 percent of its standard deviation.

Finally, with respect to state capacity, both aid modalities appear to have limited effects. While bilateral aid tends to be negatively related to state capacity, the coefficient of earmarked aid is virtually zero, and the difference in coefficients is never significant. A glance at the control variables suggests that country-level improvements in human development and governance quality are more relevant predictors of subnational governance quality. The combined take-away from the analysis so far is that, relative to the country-trend average, regions that benefit from non-delegated bilateral projects have similar development outcomes than regions without such projects, and regions with earmarked projects have slightly lower household wealth compared to regions without such projects.

**Table 2**: Aid modalities and subnational development outcomes

| | IWI | | | Child mortality | | | SCI | | |
|---|---|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) |
| Bilateral aid $_{t-2}$ | 0.083 | 0.263* | 0.119 | 0.181 | -0.450 | -0.571 | -0.252** | -0.104 | -0.051 |
| | (0.184) | (0.120) | (0.102) | (0.884) | (0.655) | (0.784) | (0.077) | (0.063) | (0.071) |
| Earmarked aid $_{t-2}$ | -0.625*** | -0.264* | -0.371** | -3.881*** | -1.136° | -0.809 | -0.119 | 0.027 | 0.014 |
| | (0.149) | (0.116) | (0.124) | (0.901) | (0.628) | (0.672) | (0.087) | (0.077) | (0.078) |
| World Bank projects $_{t-2}$ | -1.245*** | -0.310 | -0.266 | -3.085* | -0.958 | -0.686 | -0.391*** | -0.189* | -0.199° |
| | (0.309) | (0.220) | (0.267) | (1.442) | (0.892) | (1.071) | (0.100) | (0.084) | (0.113) |
| Chinese projects $_{t-2}$ | -0.015 | 0.043 | 0.012 | 0.006 | 1.304** | 1.838** | -0.219*** | -0.089 | -0.095 |
| | (0.135) | (0.089) | (0.115) | (0.683) | (0.495) | (0.602) | (0.064) | (0.055) | (0.068) |
| National HDI $_{t-2}$ | 0.313*** | 0.183*** | 0.174** | -3.405*** | -1.896*** | -2.044*** | 0.116*** | 0.096*** | 0.153*** |
| | (0.060) | (0.047) | (0.056) | (0.302) | (0.256) | (0.336) | (0.015) | (0.019) | (0.026) |
| National SCI $_{t-2}$ | 0.027 | 0.058*** | 0.021 | -0.064 | -0.375° | -0.437° | 0.690*** | 0.414*** | 0.392*** |
| | (0.025) | (0.016) | (0.018) | (0.163) | (0.207) | (0.245) | (0.011) | (0.011) | (0.014) |
| ADM1-FE | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Year-FE | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Country-period FE | No | Yes | Yes | No | Yes | Yes | No | Yes | Yes |
| PSM weights | No | No | Yes | No | No | Yes | No | No | Yes |
| Observations | 24134 | 24134 | 24134 | 22018 | 22017 | 22017 | 42316 | 42316 | 42316 |
| ADM1 clusters | 1274 | 1274 | 1274 | 1172 | 1172 | 1172 | 1671 | 1671 | 1671 |
| Within-*R2* | 0.020 | 0.005 | 0.004 | 0.056 | 0.011 | 0.013 | 0.374 | 0.114 | 0.103 |
| *p*-value – F-test | 0.004 | 0.003 | 0.003 | 0.002 | 0.460 | 0.800 | 0.263 | 0.202 | 0.552 |

*Notes*: OLS regression with various sets of high-dimensional fixed weights as shown. Dependent variables shown in column headers, and robust recipient-clustered standard errors shown in parentheses. PSM weights adjust for pre-treatment history and subnational HDI. F-test is for equality of the two aid modality coefficients. Significance levels: ° p<.1 * p<.05 ** p<.01 *** p<.001

Recognizing that not all earmarked aid is alike, I now disaggregate these aid flows into softly earmarked aid and strictly earmarked aid. I expected that softly earmarked aid would be more effective at promoting local development because it affords the implementing organizations with greater discretion as to how to use the funds. Table 3 presents the results from fixed-effects regressions for the three subnational development outcomes and different model specifications.

As regards household wealth, I find the corresponding coefficient estimates for softly earmarked aid and strictly earmarked aid to be similar. A Wald test confirms that they are not significantly different from each other. The findings confirm that bilateral aid is relatively more effective at increasing household wealth than earmarked aid, even though bilateral projects do not significantly increase wealth themselves in a consistent manner.

As concerns infant mortality, striking differences between the two types of earmarked aid emerge. While softly earmarked aid is unrelated to child mortality, strictly earmarked aid appears to significantly reduce child mortality, even in the most conservative model. In terms of effect magnitude, a region with a strictly earmarked aid project can reduce its infant mortality rate by at most 4.32 children—equivalent to about 7.9 percent of a standard deviation. Although this effect magnitude is still small, it is larger than the respective estimate for the presence of core-funded World Bank projects. In other words, donors can pressure multilaterals into effective interventions that improve child health through restrictively earmarked contributions. In fact, Wald tests confirm the difference between softly earmarked aid and strictly earmarked aid is statistically significant in all models.

Finally, turning to regional governance quality, diverging patterns between different types of earmarked aid appear to emerge as well. While coefficient estimates on softly earmarked aid tend to be positive, those for strictly earmarked aid tend to be negative, and their difference is statistically significant in the first two models. This suggests that there may be (limited) gains for governance outcomes from delegating more discretion to multilateral implementers. For bilateral aid, I obtain similar results as before, indicating no beneficial effect with respect to governance quality.

**Table 3**: Disaggregated aid modalities and subnational development outcomes

| | IWI | | | Child mortality | | | SCI | | |
|---|---|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) |
| Bilateral aid $_{t-2}$ | 0.074 | 0.260* | 0.118 | 0.128 | -0.460 | -0.569 | -0.251** | -0.104 | -0.051 |
| | (0.184) | (0.120) | (0.102) | (0.883) | (0.655) | (0.783) | (0.077) | (0.063) | (0.071) |
| Softly earmarked aid $_{t-2}$ | -0.532** | -0.383* | -0.385* | -1.659 | 0.522 | 0.796 | 0.130 | 0.186° | 0.142 |
| | (0.194) | (0.157) | (0.163) | (1.207) | (0.957) | (0.982) | (0.112) | (0.106) | (0.107) |
| Strictly earmarked aid $_{t-2}$ | -0.594*** | -0.180 | -0.327** | -4.318*** | -1.768* | -1.623* | -0.242** | -0.067 | -0.057 |
| | (0.156) | (0.115) | (0.126) | (0.973) | (0.751) | (0.791) | (0.092) | (0.081) | (0.082) |
| World Bank projects $_{t-2}$ | -1.247*** | -0.311 | -0.268 | -3.113* | -0.978 | -0.716 | -0.389*** | -0.189* | -0.200° |
| | (0.309) | (0.220) | (0.267) | (1.442) | (0.892) | (1.071) | (0.100) | (0.084) | (0.113) |
| Chinese projects $_{t-2}$ | -0.014 | 0.041 | 0.011 | 0.007 | 1.309** | 1.840** | -0.217*** | -0.087 | -0.095 |
| | (0.135) | (0.089) | (0.115) | (0.684) | (0.496) | (0.603) | (0.064) | (0.055) | (0.068) |
| National HDI $_{t-2}$ | 0.316*** | 0.185*** | 0.175** | -3.395*** | -1.897*** | -2.045*** | 0.116*** | 0.096*** | 0.153*** |
| | (0.059) | (0.047) | (0.056) | (0.302) | (0.256) | (0.336) | (0.015) | (0.019) | (0.026) |
| National SCI $_{t-2}$ | 0.026 | 0.058*** | 0.021 | -0.066 | -0.373° | -0.435° | 0.690*** | 0.414*** | 0.392*** |
| | (0.025) | (0.016) | (0.018) | (0.163) | (0.207) | (0.244) | (0.011) | (0.011) | (0.014) |
| ADM1-FE | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Year-FE | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Country-period FE | No | Yes | Yes | No | Yes | Yes | No | Yes | Yes |
| PSM weights | No | No | Yes | No | No | Yes | No | No | Yes |
| Observations | 24134 | 24134 | 24134 | 22018 | 22017 | 22017 | 42316 | 42316 | 42316 |
| ADM1 clusters | 1274 | 1274 | 1274 | 1172 | 1172 | 1172 | 1671 | 1671 | 1671 |
| Within-*R2* | 0.021 | 0.006 | 0.004 | 0.057 | 0.012 | 0.013 | 0.375 | 0.114 | 0.103 |
| Bi=So: *p*-value | 0.024 | 0.002 | 0.010 | 0.249 | 0.402 | 0.250 | 0.005 | 0.018 | 0.130 |
| Bi=Str: *p*-value | 0.008 | 0.013 | 0.006 | 0.001 | 0.196 | 0.303 | 0.939 | 0.732 | 0.961 |
| So=Str: *p*-value | 0.796 | 0.254 | 0.752 | 0.082 | 0.077 | 0.064 | 0.012 | 0.065 | 0.147 |

*Notes*: OLS regression with various sets of high-dimensional fixed weights as shown. Dependent variables shown in column headers, and robust recipient-clustered standard errors shown in parentheses. PSM weights adjust for pre-treatment history and subnational HDI. F-tests for equality of coefficients for all relevant pairs. Significance levels: ° p<.1 * p<.05 ** p<.01 *** p<.001

### 4.2. Robustness checks

I probe the robustness of these findings to alternative ways of operationalizing the key predictors. I show the full results in the appendix and report briefly on the findings here. Specifically, I replace the binary indicator for aid projects by the logged number of projects. More projects should increase the resources available for development, thereby capturing potential heterogeneity in exposure to aid projects. I obtain similar patterns as before, with increased statistical significance. Specifically, both types of earmarked aid are similarly related to decreased wealth, whereas bilateral aid appears to be relatively more effective. Strictly earmarked aid significantly reduces infant mortality, but softly earmarked aid does not. By contrast, softly earmarked aid is related to increases in subnational state capacity, whereas strictly earmarked aid is not (Table A1).

As another alternative way of measurement, I use the amounts of aid disbursed through the two funding modalities. To remove skewness, I use log-transformed aid amounts. I obtain qualitatively similar results: earmarked aid spending of all types is related to decreased household wealth, but unlike before, bilateral aid spending no longer has a positive relationship with household wealth. Strictly earmarked aid disbursements appear to reduce child mortality, whereas softly earmarked aid does not. Finally, softly earmarked aid spending no longer affects state capacity. This may be because aid amounts are unlikely to matter for governance outcomes because the primary goal of related interventions is to transfer know-how and build institutions—outcomes that are not easily quantifiable in monetary terms. Instead, measuring aid by project counts or binary variables provides a more accurate reflection of their reach and impact (Table A2).

### 4.3. Addressing inferential threats

While I have so far used already-conservative model specifications that mitigate many confounding influences, I am now deploying additional strategies to address inferential threats. In particular, I exploit variation in the timing of project decisions, mitigating selection-related confounding effects (Isaksson and Kotsadam 2018).

Table 4 broadly confirms my earlier results. With respect to the IWI, it appears that donors commit earmarked projects where adverse wealth shocks have appeared, whereas I do not observe no such selection for non-delegated bilateral projects. At the same time, the coefficients pertaining to the actual disbursement of earmarked projects are not systematically different from the coefficient of not-yet-disbursed earmarked projects. This suggests that the negative coefficient on earmarked projects obtained in previous analyses might be due to adverse selection. Moving on to child mortality, I find evidence of a negative location effect, the actual effect of project disbursement is weakly significant in only one model and insignificant in the remaining ones. Finally, with regard to state capacity, donors appear to target places with rapidly declining state capacity when deploying bilateral aid, while a similar (but weaker) selection effect appears to hold for earmarked aid. Once considering these selection effects, softly earmarked aid has a more significant positive effect on state capacity. Strictly earmarked aid and bilateral aid appear to have no effect.

**Table 4**: Accounting for not-yet-disbursed projects

| | IWI | | | Child mortality | | | SCI | | |
|---|---|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) |
| Bilateral aid $_{t-2}$ | 0.089 | 0.262* | 0.100 | 0.095 | -0.452 | -0.608 | -0.204** | -0.090 | -0.064 |
| | (0.177) | (0.119) | (0.103) | (0.859) | (0.651) | (0.775) | (0.074) | (0.063) | (0.070) |
| Committed bilateral aid | -0.156 | 0.001 | -0.125 | 0.055 | -0.197 | -0.289 | -0.298*** | -0.180*** | -0.120* |
| | (0.116) | (0.081) | (0.105) | (0.721) | (0.527) | (0.564) | (0.052) | (0.044) | (0.059) |
| Softly earmarked aid $_{t-2}$ | -0.487* | -0.374* | -0.379* | -1.484 | 0.536 | 0.801 | 0.139 | 0.185° | 0.142 |
| | (0.190) | (0.156) | (0.162) | (1.187) | (0.953) | (0.983) | (0.111) | (0.106) | (0.107) |
| Strictly earmarked aid $_{t-2}$ | -0.526*** | -0.164 | -0.313* | -4.055*** | -1.737* | -1.601* | -0.210* | -0.060 | -0.052 |
| | (0.148) | (0.113) | (0.124) | (0.922) | (0.734) | (0.781) | (0.089) | (0.080) | (0.082) |
| Committed earmarked aid | -0.713*** | -0.281** | -0.212° | -2.716** | -0.518 | -0.318 | -0.232** | -0.084 | -0.053 |
| | (0.122) | (0.097) | (0.113) | (0.853) | (0.608) | (0.697) | (0.075) | (0.070) | (0.087) |
| World Bank projects $_{t-2}$ | -1.220*** | -0.307 | -0.261 | -3.054* | -0.963 | -0.708 | -0.362*** | -0.180* | -0.193° |
| | (0.306) | (0.220) | (0.268) | (1.440) | (0.892) | (1.073) | (0.099) | (0.084) | (0.113) |
| Chinese projects $_{t-2}$ | -0.000 | 0.045 | 0.012 | 0.067 | 1.316** | 1.842** | -0.208** | -0.085 | -0.094 |
| | (0.134) | (0.089) | (0.115) | (0.688) | (0.498) | (0.605) | (0.064) | (0.055) | (0.068) |
| National HDI $_{t-2}$ | 0.319*** | 0.186*** | 0.175** | -3.381*** | -1.895*** | -2.045*** | 0.117*** | 0.094*** | 0.152*** |
| | (0.059) | (0.047) | (0.056) | (0.301) | (0.256) | (0.336) | (0.015) | (0.019) | (0.026) |
| National SCI $_{t-2}$ | 0.026 | 0.058*** | 0.022 | -0.068 | -0.373° | -0.433° | 0.689*** | 0.415*** | 0.393*** |
| | (0.025) | (0.016) | (0.018) | (0.163) | (0.207) | (0.245) | (0.011) | (0.011) | (0.014) |
| ADM1-FE | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Year-FE | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Country-period FE | No | Yes | Yes | No | Yes | Yes | No | Yes | Yes |
| PSM weights | No | No | Yes | No | No | Yes | No | No | Yes |
| Observations | 24134 | 24134 | 24134 | 22018 | 22017 | 22017 | 42316 | 42316 | 42316 |
| ADM1 clusters | 1274 | 1274 | 1274 | 1172 | 1172 | 1172 | 1671 | 1671 | 1671 |
| Within-*R2* | 0.024 | 0.006 | 0.005 | 0.058 | 0.012 | 0.013 | 0.376 | 0.115 | 0.103 |
| Bi=Com.Bi: *p*-value | 0.088 | 0.021 | 0.103 | 0.964 | 0.732 | 0.749 | 0.230 | 0.191 | 0.542 |
| So.Ear=Com.Ear: *p*-value | 0.237 | 0.530 | 0.334 | 0.310 | 0.315 | 0.372 | 0.003 | 0.019 | 0.127 |
| Str.Ear=Com.Ear: *p*-value | 0.169 | 0.291 | 0.475 | 0.103 | 0.086 | 0.166 | 0.840 | 0.808 | 0.996 |

*Notes*: OLS regression with various sets of high-dimensional fixed weights as shown. Dependent variables shown in column headers, and robust recipient-clustered standard errors shown in parentheses. PSM weights adjust for pre-treatment history and subnational HDI. F-tests for equality of coefficients for all relevant pairs. Significance levels: ° p<.1 * p<.05 ** p<.01 *** p<.001

As an alternative strategy to addressing inferential threats, I deploy shift-share instruments. As both bilateral aid and earmarked aid may be endogenous, I construct two shift-share instruments, as commonly done in the literature (Asmus-Bluhm et al. 2024; Dreher et al. 2021; Dreher and Lohmann 2015). One instrument is based on the regional probability of receiving a bilateral aid project, interacted with the annual probability of bilateral aid in the same world region. Another instrument uses the regional probability of an earmarked aid project—given the low incidence of earmarked projects that would otherwise lead to perfect separation in treatment groups—interacted with the annual probability of earmarked aid in the same world region. The instruments are strong, as reflected in their *F*-statistics well above the conventional threshold. Diagnostic tests (in the appendix) further confirm common trends across exposure groups (Figure A1) and no spurious trends in the shifter variable (Figure A2), bolstering the validity of the instrumental strategy (Christian and Barrett 2017).

The findings from instrumental-variable regressions largely support my earlier conclusions. First, I corroborate that earmarked aid is significantly negatively related to household wealth, whereas bilateral aid is insignificant. The difference in coefficients is statistically significant in the third model, which includes higher-dimensional fixed effects and observation weights. Second, I also find that earmarked aid significantly reduces child mortality, with coefficient estimates being larger than in the correlational analysis. The presence of an earmarked project reduces child mortality by half a standard deviation. In contrast, bilateral aid has the opposite effect. Third, both aid modalities also exert significantly different effects on state capacity. While bilateral aid reduces the state capacity index by about one-fifth of a standard deviation, earmarked aid has a positive effect of similar magnitude. These results are based on the models that include country-period fixed effects (Table 5).

Overall, I conclude that—for a range of inferential designs—aid modality appears to matter for aid effectiveness, considering its diverse impacts on different subnational development outcomes. Within the time horizon of two years following aid disbursements, bilateral projects do not appear to have consistent effects on household wealth but fare significantly better than earmarked projects. In contrast, (restrictively) earmarked projects significantly reduce child mortality, while bilateral aid fares significantly worse. Finally, (softly) earmarked aid appears to be most effective in building state capacity, oftentimes significantly better than strictly earmarked aid and bilateral aid.

**Table 5**: IV/2SLS analysis with shift-share instruments

| | IWI | | | Child mortality | | | SCI | | |
|---|---|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) |
| Bilateral aid | -7.187 | -3.909 | -0.449 | 59.179* | 59.494*** | 41.467*** | -0.708 | -3.870*** | -2.607*** |
| | -7.441 | -2.538 | -1.112 | -25.482 | -16.605 | -9.237 | (0.630) | (0.654) | (0.653) |
| Earmarked aid | -11.782** | -5.392*** | -5.475*** | -84.727*** | -23.706* | -31.801*** | -3.280*** | 3.003*** | 2.636** |
| | -3.863 | -1.637 | -1.212 | -16.201 | -11.353 | -8.517 | (0.759) | (0.727) | (0.833) |
| World Bank projects | 0.177 | 0.056 | 0.094 | 2.572* | 2.345*** | 2.402** | -0.100 | -0.186** | -0.167* |
| | -0.181 | -0.108 | -0.123 | -1.182 | -0.710 | -0.849 | (0.069) | (0.064) | (0.078) |
| Chinese projects | -0.480 | -0.092 | -0.260 | -9.396** | -6.529*** | -4.778** | -0.196 | 0.111 | -0.064 |
| | -0.846 | -0.300 | -0.290 | -3.422 | -1.863 | -1.740 | (0.132) | (0.110) | (0.134) |
| National HDI | 0.402*** | 0.228*** | 0.215*** | -2.215*** | -1.829*** | -1.597*** | 0.152*** | 0.084*** | 0.128*** |
| | -0.071 | -0.051 | -0.060 | -0.412 | -0.309 | -0.414 | (0.016) | (0.021) | (0.028) |
| National SCI | 0.035 | 0.066*** | 0.028 | -0.197 | -0.379° | -0.456° | 0.680*** | 0.407*** | 0.387*** |
| | -0.032 | -0.018 | -0.020 | -0.203 | -0.220 | -0.269 | (0.012) | (0.012) | (0.015) |
| ADM1-FE | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Year-FE | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Country-period FE | No | Yes | Yes | No | Yes | Yes | No | Yes | Yes |
| PSM weights | No | No | Yes | No | No | Yes | No | No | Yes |
| Observations | 24134 | 24134 | 24134 | 22018 | 22018 | 22018 | 42316 | 42316 | 42316 |
| ADM1 clusters | 1274 | 1274 | 1274 | 1172 | 1172 | 1172 | 1671 | 1671 | 1671 |
| Adjusted *R2* | 0.428 | 0.750 | 0.771 | -0.587 | 0.234 | 0.283 | 0.301 | 0.354 | 0.357 |
| Kleibergen-Paap F-stat | 7.322 | 16.774 | 41.419 | 20.262 | 18.437 | 38.935 | 244.399 | 166.815 | 126.605 |
| Bi=Ear: *p*-value | 0.325 | 0.463 | 0.003 | 0.000 | 0.000 | 0.000 | 0.026 | 0.000 | 0.000 |

*Notes*: IV/2SLS regression with various sets of high-dimensional fixed effects and weights as shown. Dependent variables shown in column headers, and robust recipient-clustered standard errors shown in parentheses. PSM weights adjust for pre-treatment history and subnational HDI. F-tests for equality of coefficients among aid modalities. Significance levels: ° p<.1 * p<.05 ** p<.01 *** p<.001

## 5. Conclusion

This paper examined the effectiveness of aid projects with different implementation modalities, specifically whether or not a bilateral project is delegated to a multilateral organization. Following principal-agent theory, which highlights the unique expertise of multilateral bureaucracies, I expected that delegation is related to increased effectiveness, compared to bilateral aid without multilateral delegation. However, as a growing literature on earmarked funding has shown, multilaterals have limited capacity to effectively deliver donor-driven earmarked projects, given the lack of flexibility that these projects entail and the additional reporting requirements that donors impose upon multilaterals. Whether the delegation benefits of earmarked projects outweigh their costs is an open question and therefore how earmarked projects perform vis-à-vis bilateral projects is an empirical question.

I combined data on the geographic location of aid projects of 19 bilateral DAC donors and the funding modalities of these projects to construct a database for 1,753 administrative regions (ADM1-level) from 1990 to 2022. Using different estimation strategies, I find generally no significant relationship between aid projects and subnational development outcomes. The exception is a positively significant effect of earmarked projects on the human development index using instrumental-variable analysis. Yet, none of the conventional regression estimates, with various sets of fixed effects and pre-treatment matching, yield any significant relationship between aid projects and development outcomes in the DAC aggregate. This is not to say that aid projects never have any effects: I found differences across donors and aid modalities, as well as when distinguishing levels of earmarking stringency.

Before elaborating the broader implications of the findings, I discuss the limitations. First, while geolocated data holds promise to enhance our inferential leverage, a large proportion of earmarked projects is not allocable to a specific location. In fact, earmarked projects can support institutional reform initiatives at headquarter level which can enhance the productivity of all projects alike. As I excluded these projects from the analysis, my estimates might represent a lower bound of the true effect. Second, my estimation setup focused on identifying short-term effects, considering the period during which a project disburses and the two years after project termination. In reality, projects might take longer to yield demonstrable effects. Finally, I focused on aggregate effects for different funding modalities, neglecting possible differences across aid sectors as well as differences over time. Heterogeneity analysis might reveal that effects differ in certain sub-samples.

As the first analysis of the effectiveness of earmarked projects against bilateral projects, my paper makes several contributions. First and foremost, the findings close an important gap in our understanding of the effectiveness of earmarked funding (Ihl et al. 2025). Previous research compared earmarked projects to core-funded projects, assuming that core-funded projects are the natural benchmark (Heinzel et al. 2024a; Heinzel and Reinsberg 2024). While research found earmarked projects to be less effective than core-funded projects, it was unknown how these projects performed relative to bilateral projects not delegated to a multilateral organization. This paper is the first to fill this important gap, allowing donors contemplating about implementation modalities to make more informed choices for delivering their aid programs. Methodologically, my paper addresses a limitation in the existing research on earmarked funding which must assume that core-funded projects are in fact like earmarked projects. In some cases, bilateral projects that could be delivered by a range of different implementers in the marketplace for development might provide a better pool of comparable projects.

For the debate about aid effectiveness, my paper extends the growing evidence base on aid effectiveness at the subnational level (Bitzer and Gören 2024; Cruzatti et al. 2023; Dreher et al. 2021), emphasizing the hitherto overlooked role of funding modalities. In that regard, I echo long-heeded views in the cross-country aid literature that delivery modalities matter (Dietrich 2021; Ouattara and Strobl 2008; Wright and Winters 2010). At the same time, I do not find any delivery modality to be superior but instead that the costs and benefits of different aid modalities are highly continent and involve tradeoffs. While previous aid literature has already highlighted several tradeoffs between aid modalities (Michaelowa et al. 2018; Milner and Tingley 2013; Reinsberg et al. 2017), my work highlights new dilemmas. Specifically, I showed that bilateral aid might be the preferred option for boosting incomes, but it does poorly with respect to building more capable states. Earmarked aid, by contrast, involves its own tradeoffs. Restricting how earmarked aid can be used might lower child mortality, but it could also hinder the development of effective state institutions. A better approach for donors might be to only softly earmark their contributions.

Overall, the results highlight the need for caution when trying to make recommendations on how donors should engage their bilateral aid budgets. On the one hand, my results cast doubt on the moniker that earmarked funding combines 'the best of two worlds'—as its benefits appear more contingent with respect to both the type of development impacts that should be reached and the concrete design of earmarked funding modalities. In other words, the purported benefits of delegation that proponents of earmarked aid have often highlighted appear to be overblown. On the other hand, the results mitigate concerns that earmarked aid undermines aid effectiveness. Even tightly earmarked aid can be beneficial for a subset of development outcomes. This does not invalidate concerns that to the extent that donor use multilaterals as implementers of bilateral priorities, they quickly deplete their multilateral assets. As an increasing number of donors (tightly) earmark their contributions, the capacity of multilaterals becomes strained and their performance declines. Yet, multilaterals work hard to cope with these pressures, given that the performance losses induced through tight delegation do not seem to systematically affect aid impact. That said, donors need to develop a clearer understanding of when using multilaterals is appropriate and when their goals may be better served with bilateral aid. This could provide the multilateral system with the breathing space that it needs to focus on its core missions, from alleviating poverty to providing global public goods.

## References


Abbott, K. W., & Snidal, D. (1998). Why states act through formal international organizations. *Journal of Conflict Resolution*, *42*(1), 3–32. https://doi.org/10.1177/0022002798042001001

Acharya, Arnab, Fuzzo de Lima, A. T., & Moore, M. (2006). Proliferation and fragmentation: Transactions costs and the value of aid. *Journal of Development Studies*, *42*(1), 1–21.

Ahlfeldt, G. M., Nitsch, V., & Wendland, N. (2019). Ease vs. noise: Long-run changes in the value of transport (dis) amenities. *Journal of Environmental Economics and Management*, *98*(102268).

Annen, K., & Knack, S. (2015). *On the Delegation of Aid Implementation to Multilateral Agencies*. Washington D.C.

Ashton, L., Friedman, J., Goldemberg, D., Hussain, M. Z., Kenyon, T., Khan, A., & Zhou, M. (2023). A puzzle with missing pieces: Explaining the effectiveness of World Bank development projects. *World Bank Research Observer*, *38*(1), 115–146.

Asmus-Bluhm, G., Eichenauer, V. Z., Fuchs, A., & Parks, B. (2024). Does India Use Development Finance to Compete With China? A Subnational Analysis. *Journal of Conflict Resolution*, *69*(2–3), 00220027241228184.

Barakat, S., Rzeszut, K., & Martin, N. (2012). *What is the track record of Multi-Donor Trust Funds in improving aid effectiveness?*

Barnett, M. N., & Finnemore, M. (2004). *Rules for the world: International organizations in global politics*. Ithaca: Cornell University Press.

Bauer, M. W., & Ege, J. (2016). Bureaucratic autonomy of international organizations' secretariats. *Journal of European Public Policy*, *23*(7), 1019–1037.

Baumann, M.-O., Lundsgaarde, E., & Weinlich, S. (2019). Shades of grey: Earmarking in the UN development system. In *Financing the UN Development System: Time for hard choices* (pp. 106–109). New York: UNDP.

Baumann, M.-O., & Weinlich, S. (2020). Funding the UN: Support or constraint? In S. Browne & T. G. Weiss (Eds.), *Routledge Handbook on the UN and Development Funding the UN Support or constraint?* (pp. 151–164). Taylor & Francis Group.

Bayram, A. B., & Graham, E. R. (2017). Financing the United Nations: Explaining variation in how donors provide funding to the UN. *Review of International Organizations*, *12*, 421–459. https://doi.org/10.1007/s11558-016-9261-0

Bitzer, J., Dannemann, B. C., & Gören, E. (2025). *Foreign aid and local conflict dynamics: A monthly grid-cell-level analysis in Africa* (No. 425). Oldenburg.

Bitzer, J., & Gören, E. (2024). Foreign aid and subnational development: A grid cell analysis. *Journal of Development Studies*, *60*(10), 1557–1591.

Bluhm, R., Dreher, A., Fuchs, A., Parks, B. C., Strange, A. M., & Tierney, M. J. (2025). Connective financing: Chinese infrastructure projects and the diffusion of economic activity in developing countries. *Journal of Urban Economics*, *145*(103730).

Bomprezzi, P., Dreher, A., Fuchs, A., Hailer, T., Kammerlander, A., Kaplan, L., et al. (2024). Wedded to Prosperity? Informal Influence and Regional Favoritism. *CEPR Discussion Paper*, *18878*.

Borusyak, K., Hull, P., & Jaravel, X. (2022). Quasi-Experimental Shift-Share Research Designs. *Review of Economic Studies*, *89*(1), 181–213. https://doi.org/10.1093/restud/rdab030

Brazys, S., Elkink, J. A., & Kelly, G. (2017). Bad Neighbors? How Co-Located Chinese and World Bank Development Projects Impact Local Corruption in Tanzania. *Review of International Organizations*, *12*(2), 227–253.

Bueno de Mesquita, B., & Smith, A. (2009). A Political Economy of Aid. *International Organization*, *63*(2), 309–340.

Bulman, D., Kolkma, W., & Kraay, A. (2017). Good countries or good projects? Comparing macro and micro correlates of World Bank and Asian Development Bank project performance. *Review of International Organizations*, *12*(3), 335–363. https://doi.org/10.1007/s11558-016-9256-x

Campbell, S. (2018). *Global Governance and Local Peace: Accountability and Performance in International Peacebuilding*. New York: Cambridge University Press.

Chauvet, L., Collier, P., & Duponchel, M. (2010). What explains aid project success in post-conflict situations? *World Bank Policy Research Working Paper*, (5418).

Cheeseman, N., Swedlund, H. J., & O'Brien-Udry, C. (2024). Foreign aid withdrawals and suspensions: Why, when and are they effective? *World Development*, *178*(106571).

Christian, P., & Barrett, C. B. (2017). *Revisiting the effect of food aid on conflict: A methodological caution*. Washington D.C.

Crombach, L., & Smits, J. (2024). The subnational corruption database: Grand and petty corruption in 1,473 regions of 178 countries, 1995–2022. *Scientific Data*, *11*(1), 686.

Cruzatti, J., Dreher, A., & Matzat, J. (2023). Chinese aid and health at the country and local level. *World Development*, *167*(106214).

Custer, S., Dreher, A., Elston, T., Fuchs, A., Ghose, S., Lin, J. J., et al. (2021). *Tracking Chinese Development Finance: An Application of AidData's TUFF 2.0 Methodology*. *AidData Working Paper*.

Daoud, A., & Reinsberg, B. (2019). Structural adjustment, state capacity and child health: Evidence from IMF programmes. *International Journal of Epidemiology*, *48*(2), 445–454. https://doi.org/10.1093/ije/dyy251

Dietrich, S. (2021). *States, Markets, and Foreign Aid*. Cambridge: Cambridge University Press.

Doucouliagos, H., & Paldam, M. (2009). The aid effectiveness literature: The sad results of 40 years of research. *Journal of Economic Surveys*, *23*(3), 433–461. https://doi.org/10.1111/j.1467-6419.2008.00568.x

Dreher, A., Fuchs, A., Hodler, R., Parks, B. C., Raschky, P. A., & Tierney, M. J. (2021). Is Favoritism a Threat to Chinese Aid Effectiveness? A Subnational Analysis of Chinese Development Projects. *World Development*, *105291*.

Dreher, A., & Kilby, C. (2010). The impact of aid on growth revisited: Do donor motives matter? *Economic Letters*, *107*(3), 338–340.

Dreher, A., Klasen, S., Vreeland, J. R., & Werker, E. (2013). The Costs of Favoritism: Is Politically Driven Aid Less Effective? *Economic Development and Cultural Change*, *62*(1), 157–191. https://doi.org/10.1086/671711

Dreher, A., Lang, V. F., & Reinsberg, B. (2024). Aid effectiveness and donor motives. *World Development*, *176*(106501).

Dreher, A., Lang, V., Rosendorff, B. P., & Vreeland, J. R. (2021). Bilateral or Multilateral? International Financial Flows and the Dirty-Work Hypothesis. *Journal of Politics*.

Dreher, A., & Lohmann, S. (2015). Aid and growth at the regional level. *Oxford Review of Economic Policy*, *31*(3), 420–446.

Dreher, A., Nunnenkamp, P., & Schmaljohann, M. (2015). The allocation of German aid: Self-interest and government ideology. *Economics & Politics*, *27*(1), 160–184.

Dreher, A., Pan, J., Schneider, C., & Tang, K. (2025). The Electoral Effects of Foreign Aid: Evidence from Sub-National Data. In *Paper presented at "Analyzing Global Development with the Geocoded Official Development Assistance Dataset (GODAD)"*.

Dunning, T. (2004). Conditioning the Effects of Aid: Cold War Politics, Donor Credibility, and Democracy in Africa. *International Organization*, *58*, 409–423.

Easterly, W., & Pfutze, T. (2008). Where Does the Money Go? Best and Worst Practices in Foreign Aid. *Journal of Economic Perspectives*, *22*(2), 29–52. https://doi.org/10.1257/jep.22.2.29

Eckhard, S., & Parizek, M. (2022). Policy Implementation by International Organizations: A Comparative Analysis of Strengths and Weaknesses of National and International Staff. *Journal of Comparative Policy Analysis: Research and Practice*, *24*(3), 254–270. https://doi.org/10.1080/13876988.2020.1813032

Eichenauer, V. Z., & Reinsberg, B. (2017). What determines earmarked funding to international development organizations? Evidence from the new multi-bi aid data. *Review of International Organizations*, *12*(2), 171–197. https://doi.org/10.1007/s11558-017-9267-2

Eilers, Y., Kluve, J., Langbein, J., & Reiners, L. (2025). Volume, Risk, Complexity: What Makes Development Finance Projects Succeed or Fail? *World Bank Economic Review*, https://doi.org10.1093/wber/lhaf001. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4682475

Fidler, D. P. (2016). Health. In Amitav Acharya (Ed.), *Why govern? Rethinking Demand and Progress in Global Governance* (p. chapter 12). Cambridge: Cambridge University Press.

Frey, B. S. (1997). The Public Choice of International Organizations. *Mueller, Dennis (ed.). Perspectives on Public Choice.*

Gatti, F. (2024). Contesting focality in global health governance: The World Health Organization-World Bank relationship in historical perspective. *British Journal of Politics and International Relations*, doi: 10.1177/13691481241293063.

Gehring, K., Kaplan, L. C., & Wong, M. H. (2022). China and the World Bank—How contrasting development approaches affect the stability of African states. *Journal of Development Economics*, *158*(102902).

Gisselquist, R. M., & Tarp, F. (2019). Aid Impact and Effectiveness: Introduction and Overview. *Politics and Governance*, *7*(2), 1. https://doi.org/10.17645/pag.v7i2.2219

Graham, E. R. (2023). *Transforming International Institutions: How Money Quietly Sidelined Multilateralism at The United Nations*. Oxford: Oxford University Press.

Hainmueller, J. (2012). Entropy balancing for causal effects: A multivariate reweighting method to produce balanced samples in observational studies. *Political Analysis*, *20*(1), 25–46. https://doi.org/10.1093/pan/mpr025

Hawkins, D. G., Lake, D. A., Nielson, D. L., & Tierney, M. J. (2006). Delegation under anarchy: states, international organizations, and principal-agent theory. In *Delegation and Agency in International Organizations* (pp. 3–38). Cambridge: Cambridge University Press.

Headey, D. (2008). Geopolitics and the effect of foreign aid on economic growth: 1970–2001. *Journal of International Development*, *20*(2), 161–180.

Heinzel, M. (2022). International Bureaucrats and Organizational Performance: Country-Specific Knowledge and Sectoral Knowledge in World Bank Projects. *International Studies Quarterly*, *66*(2), sqac013.

Heinzel, M., Cormier, B., & Reinsberg, B. (2023). Earmarked funding and the control−performance tradeoff in international development organizations. *International Organization*, *77*(2), 475–495.

Heinzel, M., & Reinsberg, B. (2024). Trust funds and the sub-national effectiveness of development aid: Evidence from the World Bank. *World Development*, *179*(106609), doi: 10.1016/j.worlddev.2024.106609.

Heinzel, M., Reinsberg, B., & Zaccaria, G. (2024a). Core funding and the performance of

international organizations: Evidence from UNDP projects. *Regulation & Governance*, doi: 10.1111/rego.12632.

Heinzel, M., Reinsberg, B., & Zaccaria, G. (2024b). Issue congruence in international organizations: A study of World Bank spending. *Global Policy*, doi: 10.1111/1758-5899.13413.

Honig, D. (2019). When Reporting Undermines Performance: The Costs of Politically Constrained Organizational Autonomy in Foreign Aid Implementation. *International Organization*, *73*(1), 171–201.

Ihl, J., Singh, R., Mukali, I. M., Skagerlind, H. H., & Heucher, A. (2025). *The effectiveness of core and earmarked funding in multilateral development cooperation* (No. DEval Discussion Papers 02/2025). Bonn.

Isaksson, A. S., & Kotsadam, A. (2018). Racing to the bottom? Chinese development projects and trade union involvement in Africa. *World Development*, *106*, 284–298. https://doi.org/10.1016/j.worlddev.2018.02.003

Kaya, A., Kilby, C., & Kay, J. (2021). Asian Infrastructure Investment Bank as an instrument for Chinese influence? Supplementary versus remedial multilateralism. *World Development*, *145*, 105531. https://doi.org/10.1016/j.worlddev.2021.105531

Kilby, C. (2013). An Empirical Assessment of Informal Influence in the World Bank. *Economic Development and Cultural Change*, *61*(2), 431–464. https://doi.org/10.1086/668278

Kilby, C. (2015). Assessing the impact of World Bank preparation on project outcomes. *Journal of Development Economics*, *115*, 111–123.

Knack, S., & Rahman, A. (2007). Donor fragmentation and bureaucratic quality in aid recipients. *Journal of Development Economics*, *83*(1), 176–197.

Kotsadam, A., Østby, G., Rustad, S. A., Tollefsen, A. F., & Urdal, H. (2018). Development aid and infant mortality: Micro-level evidence from Nigeria. *World Development*, *105*, 59–69.

Lall, R. (2017). Beyond Institutional Design: Explaining the Performance of International Organizations. *International Organization*, *71*(2), 245–280. https://doi.org/10.1017/S0020818317000066

Lang, V. F. (2020). The Economics of the Democratic Deficit: The Effect of IMF Programs on Inequality. *Review of International Organizations*, forthcoming. https://doi.org/10.2139/ssrn.2883120

Licht, A. A. (2010). Coming into Money: The Impact of Foreign Aid on Leader Survival. *Journal of Conflict Resolution*, *54*(1), 58–87. https://doi.org/10.1177/0022002709351104

Lipson, M. (2010). Performance under ambiguity: International organization performance in UN peacekeeping. *Review of International Organizations*, *5*(1), 249–284.

Martens, B., Mummert, U., Murrell, P., & Seabright, P. (2003). *The Institutional Economics of Foreign Aid*. https://doi.org/10.1016/S0304-3878(03)00041-5

Martorano, B., Metzger, L., & Sanfilippo, M. (2020). Chinese Development Assistance and Household Welfare in Sub-Saharan Africa. *World Development*, *129*(104909).

Michaelowa, K., Reinsberg, B., & Schneider, C. (2018). The Politics of Delegation in the European Union. *International Studies Quarterly*. https://doi.org/10.1093/isq/sqy034

Milner, H. V, & Tingley, D. (2013). The choice for multilateralism: Foreign aid and American

foreign policy. *Review of International Organizations*, *8*, 313–341.

Minoiu, C., & Reddy, S. G. (2010). Development aid and economic growth: A positive long-run relation. *Quarterly Review of Economics and Finance*, *50*(1), 27–39.

Moravcsik, A. (2004). Is there a "Democratic Deficit" in World Politics? A Framework for Analysis. *Government and Opposition*, *39*(2), 19–36. https://doi.org/10.1111/j.1477-7053.2004.00126.x

Nielson, D. L., & Tierney, M. J. (2003). Delegation to International Organizations: Agency Theory and World Bank Environmental Reform. *International Organization*, *57*(02), 241–276.

Niskanen, W. A. (1971). Bureaucracy and Representative Government. Transaction Publishers. https://books.google.co.uk/books?id=dOYe1ld9F1QC

OECD. (2011). *2011 DAC Report on Multilateral Aid*. Paris. DCD/DAC(2011)21/FINAL

OECD. (2022). Aggregate official and private flows (DAC1a). Paris: Organization for Economic Cooperation and Development (OECD). stats.oecd.org

Ouattara, B., & Strobl, E. (2008). Aid, policy and growth: Does aid modality matter? *Review of World Economics*, *144*(2), 347–365. https://doi.org/10.1007/s10290-008-0150-3

Radelet, S. (2006). *A Primer on Foreign Aid*. *Center for Global Development Working Paper*. papers2://publication/uuid/7F92956E-A7DC-4640-9EAC-2A20B28CAB39

Ravallion, M. (2016). The World Bank: Why It Is Still Needed and Why It Still Disappoints. *The Journal Of Economic Perspectives*, *30*(1), 77–94. https://doi.org/10.1257/jep.30.1.77

Reinsberg, B. (2023). Earmarked funding and the performance of international organizations: Evidence from food and agricultural development agencies. *Global Studies Quarterly*, *3*(4), ksad056.

Reinsberg, B., Heinzel, M., & Siauwijaya, C. (2023). Tracking earmarked funding to international organizations: Introducing the Earmarked Funding Dataset. Glasgow: University of Glasgow. https://doi.org/10.7910/DVN/WCS99H

Reinsberg, B., Heinzel, M., & Siauwijaya, C. (2024). Tracking earmarked funding in international organizations: Introducing the Earmarked Funding Dataset. *Review of International Organizations*, doi: 10.1007/s11558-024-09548-1.

Reinsberg, B., Michaelowa, K., & Knack, S. (2017). Which Donors, Which Funds? Bilateral Donors' Choice of Multilateral Funds at the World Bank. *International Organization*, *71*(4), 767–802. https://doi.org/10.1017/S0020818317000340

Reinsberg, B., & Siauwijaya, C. (2024). Does earmarked funding affect the performance of international organizations? *Global Policy*, *15*(1), 23–39.

Reinsberg, B., & Taggart, J. R. (2025). Does earmarked foreign aid undermine recipient-country ownership? *Journal of International Development*, doi: 10.1002/jid.3985).

Rodrik, D. (1995). *Why is there multilateral lending?*

Rustad, S. A., Hoelscher, K., Kotsadam, A., Ostby, G., & Urdal, H. (2019). *Does Development Aid Address Political Exclusion? A Disaggregated Study of the Location of Aid in Sub-Saharan Africa*.

Schmid, L., Reitzenstein, A., & Hall, N. (2021). Blessing or Curse? The Effects of Earmarked Funding in UNICEF and UNDP. *Global Governance*, *27*(1), 433–459.

Smits, J. (2016). *GDL Area Database: Sub-national development indicators for research and policy making.* Nijmegen.

Smits, J., & Permanyer, I. (2019). The subnational human development database. *Scientific Data*, *6*(1), 1–15.

Smits, J., & Steendijk, R. (2015). The international wealth index (IWI). *Social Indicators Research*, *122*(1), 65–85.

Sridhar, D., & Woods, N. (2013). Trojan Multilateralism: Global Cooperation in Health. *Global Policy*, *4*(4), 325–335.

Stone, R. W. (2011). *Controlling Institutions: international organizations and the global economy*. Cambridge: Cambridge University Press.

Swedlund, H. J. (2017). Can foreign aid donors credibly threaten to suspend aid? Evidence from a cross-national survey of donor officials. *Review of International Political Economy*, *24*(3), 454–496.

Tortora, P., & Steensen, S. (2014). *Making earmarked funding more effective: Current practices and a way forward*. Paris: OECD. https://www.oecd.org/dac/aid-architecture/Multilateral Report N 1_2014.pdf

UN-MPTFO. (2022). *Financing the UN Development System: Joint Responsibilities in a World of Disarray*. New York. https://financingun.report/sites/default/files/2022-09/DHF-Financing-the-UN-Development-System-2022.pdf

UN. (2019). *Implementation of General Assembly resolution 71/243 on the quadrennial comprehensive policy review of operational activities for development of the United Nations system, 2019: funding compact*. New York.

Vaubel, R. (2006). Principal-agent problems in international organizations. *The Review of International Organizations*, *1*(2), 125–138.

Weaver, C. E. (2008). *The World Bank's Hypocrisy and the Poverty of Reform*. Ithaca: Cornell University Press.

Weinlich, S., Baumann, M.-O., Lundsgaarde, E., & Wolff, P. (2020). *Earmarking in the multilateral developoment system: Many shades of grey*. Bonn.

Wellner, L., Dreher, A., Fuchs, A., Parks, B. C., & Strange, A. M. (2025). Can Aid Buy Foreign Public Support? Evidence from Chinese Development Finance. *Economic Development and Cultural Change*, *73*(2), 523–578.

Winters, M. S. (2019). Too many cooks in the kitchen? The division of financing in World Bank projects and project performance. *Politics and Governance*, *7*(2), 117–126. https://doi.org/10.17645/pag.v7i2.1826

Wright, J., & Winters, M. S. (2010). The politics of effective foreign aid. *Annual Review of Political Science*, *13*, 61–80.

## Supplemental appendix

**Table A1**: Number of aid projects and subnational development outcomes

| | IWI | | | Child mortality | | | SCI | | |
|---|---|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) |
| # Bilateral projects | 0.177 | 0.052 | -0.010 | -1.343* | -0.474 | -0.109 | -0.176*** | -0.103* | -0.056 |
| | (0.116) | (0.076) | (0.077) | (0.594) | (0.447) | (0.500) | (0.049) | (0.044) | (0.048) |
| # Softly earmarked projects | -0.534** | -0.350* | -0.341* | -1.900 | 0.242 | 0.581 | 0.233° | 0.241* | 0.200° |
| | (0.200) | (0.154) | (0.162) | (1.457) | (1.149) | (1.134) | (0.125) | (0.115) | (0.117) |
| # Strictly earmarked projects | -0.674*** | -0.265* | -0.379** | -4.672*** | -1.812* | -1.815° | -0.213* | -0.074 | -0.065 |
| | (0.167) | (0.127) | (0.145) | (1.088) | (0.844) | (0.927) | (0.090) | (0.083) | (0.087) |
| World Bank projects | -1.256*** | -0.297 | -0.261 | -3.110* | -1.003 | -0.765 | -0.382*** | -0.185* | -0.195° |
| | (0.313) | (0.220) | (0.267) | (1.447) | (0.892) | (1.069) | (0.101) | (0.084) | (0.114) |
| Chinese projects | -0.020 | 0.037 | 0.011 | 0.132 | 1.330** | 1.849** | -0.194** | -0.078 | -0.090 |
| | (0.134) | (0.089) | (0.115) | (0.680) | (0.493) | (0.602) | (0.064) | (0.055) | (0.068) |
| National HDI | 0.314*** | 0.185*** | 0.175** | -3.358*** | -1.884*** | -2.036*** | 0.123*** | 0.098*** | 0.153*** |
| | (0.059) | (0.047) | (0.056) | (0.301) | (0.256) | (0.338) | (0.015) | (0.019) | (0.026) |
| National SCI | 0.027 | 0.059*** | 0.022 | -0.074 | -0.373° | -0.435° | 0.687*** | 0.413*** | 0.392*** |
| | (0.025) | (0.016) | (0.018) | (0.163) | (0.207) | (0.246) | (0.011) | (0.011) | (0.014) |
| ADM1-FE | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Year-FE | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Country-period FE | No | Yes | Yes | No | Yes | Yes | No | Yes | Yes |
| PSM weights | No | No | Yes | No | No | Yes | No | No | Yes |
| Observations | 24134 | 24134 | 24134 | 22018 | 22017 | 22017 | 42316 | 42316 | 42316 |
| ADM1 clusters | 1274 | 1274 | 1274 | 1172 | 1172 | 1172 | 1671 | 1671 | 1671 |
| Within-*R2* | 0.022 | 0.005 | 0.004 | 0.059 | 0.012 | 0.013 | 0.375 | 0.115 | 0.103 |
| Bi=So: *p*-value | 0.002 | 0.019 | 0.062 | 0.728 | 0.575 | 0.588 | 0.004 | 0.007 | 0.051 |
| Bi=Str: *p*-value | 0.000 | 0.033 | 0.022 | 0.009 | 0.188 | 0.124 | 0.720 | 0.767 | 0.928 |
| So=Str: *p*-value | 0.586 | 0.643 | 0.844 | 0.125 | 0.164 | 0.112 | 0.006 | 0.038 | 0.088 |

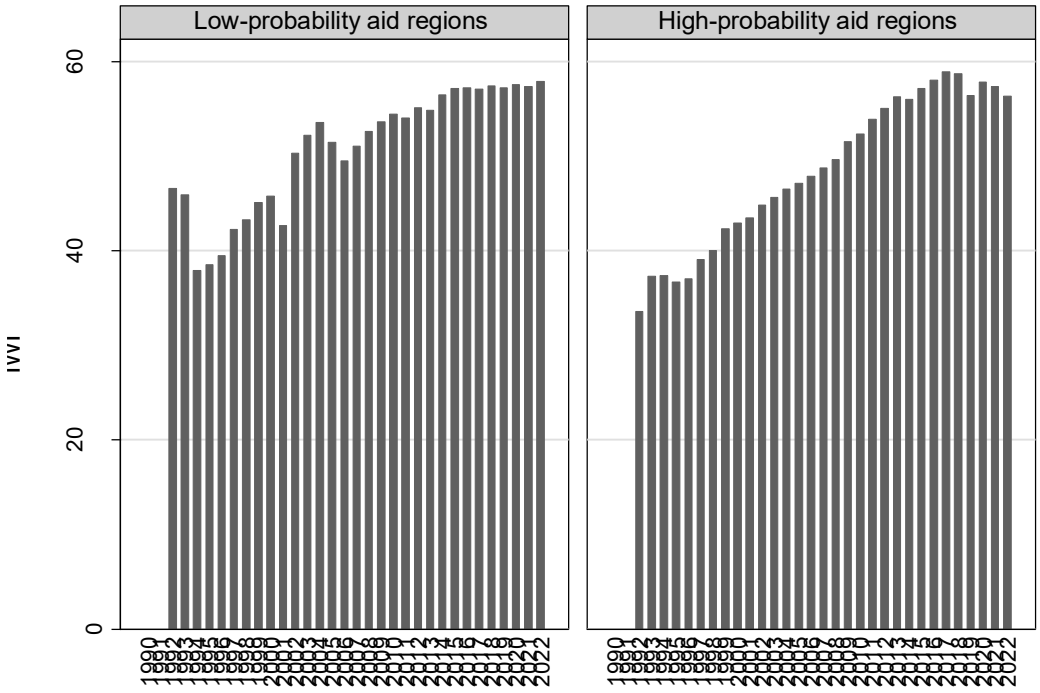
*Notes*: OLS regression with various sets of high-dimensional fixed weights as shown. Dependent variables shown in column headers, and robust recipient-clustered standard errors shown in parentheses. PSM weights adjust for pre-treatment history and subnational HDI. F-tests for equality of coefficients for all relevant pairs. Significance levels: ° p<.1 * p<.05 ** p<.01 *** p<.001

**Table A2**: Logged aid amounts and subnational development outcomes

| | IWI | | | Child mortality | | | SCI | | |
|---|---|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) |
| Bilateral US$ | -0.075*** | -0.024° | -0.018 | -0.077 | -0.025 | -0.058 | -0.044*** | -0.016** | -0.006 |
| | (0.020) | (0.013) | (0.016) | (0.097) | (0.059) | (0.067) | (0.007) | (0.006) | (0.009) |
| Softly earmarked US$ | -0.043** | -0.031* | -0.031* | -0.186° | 0.028 | 0.058 | 0.008 | 0.012 | 0.009 |
| | (0.017) | (0.013) | (0.013) | (0.107) | (0.080) | (0.082) | (0.010) | (0.009) | (0.009) |
| Strictly earmarked US$ | -0.057*** | -0.018° | -0.029* | -0.366*** | -0.132* | -0.123° | -0.027*** | -0.010 | -0.009 |
| | (0.013) | (0.011) | (0.012) | (0.084) | (0.065) | (0.067) | (0.008) | (0.007) | (0.007) |
| World Bank projects | -0.715* | -0.141 | -0.142 | -2.614° | -0.868 | -0.407 | -0.040 | -0.077 | -0.158 |
| | (0.296) | (0.208) | (0.259) | (1.479) | (0.907) | (1.080) | (0.108) | (0.091) | (0.121) |
| Chinese projects | -0.005 | 0.039 | 0.011 | 0.050 | 1.316** | 1.838** | -0.204** | -0.084 | -0.094 |
| | (0.134) | (0.089) | (0.115) | (0.683) | (0.495) | (0.602) | (0.064) | (0.055) | (0.068) |
| National HDI | 0.321*** | 0.186*** | 0.176** | -3.384*** | -1.894*** | -2.033*** | 0.119*** | 0.097*** | 0.153*** |
| | (0.059) | (0.047) | (0.056) | (0.302) | (0.256) | (0.337) | (0.015) | (0.019) | (0.026) |
| National SCI | 0.022 | 0.057*** | 0.021 | -0.076 | -0.376° | -0.439° | 0.687*** | 0.414*** | 0.392*** |
| | (0.025) | (0.016) | (0.018) | (0.164) | (0.207) | (0.246) | (0.011) | (0.011) | (0.014) |
| ADM1-FE | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Year-FE | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Country-period FE | No | Yes | Yes | No | Yes | Yes | No | Yes | Yes |
| PSM weights | No | No | Yes | No | No | Yes | No | No | Yes |
| Observations | 24134 | 24134 | 24134 | 22018 | 22017 | 22017 | 42316 | 42316 | 42316 |
| ADM1 clusters | 1274 | 1274 | 1274 | 1172 | 1172 | 1172 | 1671 | 1671 | 1671 |
| Within-*R2* | 0.024 | 0.006 | 0.004 | 0.057 | 0.011 | 0.013 | 0.376 | 0.115 | 0.103 |
| Bi=So: *p*-value | 0.055 | 0.727 | 0.931 | 0.373 | 0.861 | 0.631 | 0.000 | 0.000 | 0.019 |
| Bi=Str: *p*-value | 0.205 | 0.603 | 0.612 | 0.016 | 0.182 | 0.440 | 0.003 | 0.059 | 0.610 |
| So=Str: *p*-value | 0.314 | 0.853 | 0.596 | 0.110 | 0.311 | 0.210 | 0.001 | 0.022 | 0.023 |

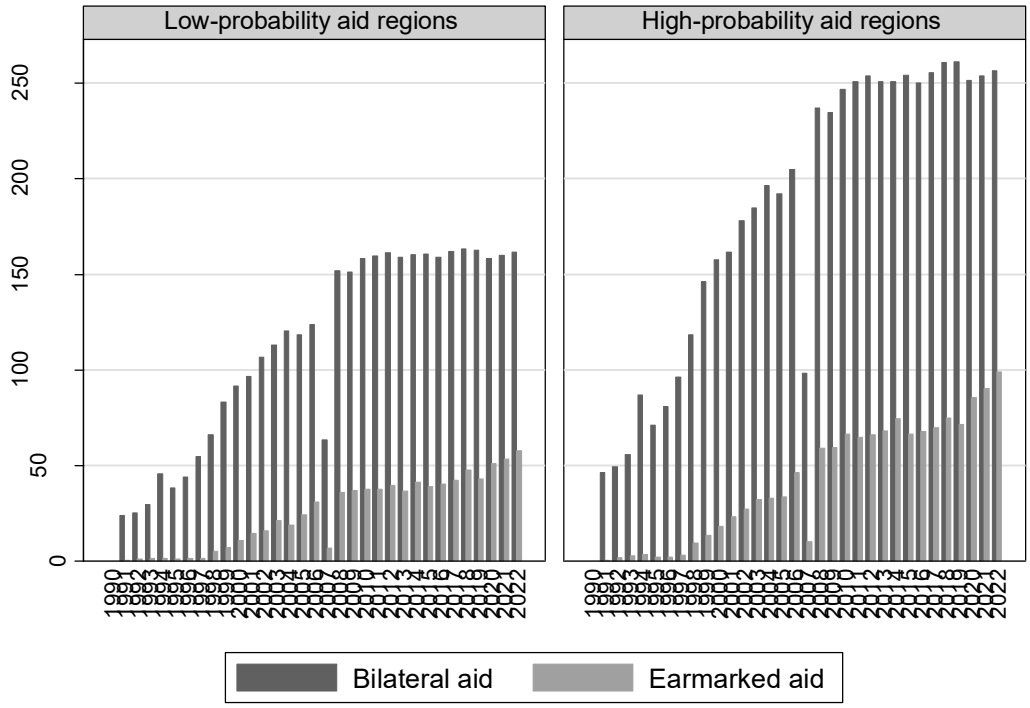
*Notes*: OLS regression with various sets of high-dimensional fixed weights as shown. Dependent variables shown in column headers, and robust recipient-clustered standard errors shown in parentheses. PSM weights adjust for pre-treatment history and subnational HDI. F-tests for equality of coefficients for all relevant pairs. Significance levels: ° p<.1 * p<.05 ** p<.01 *** p<.001

**Figure A1**: Trends in outcome variables across exposure groups



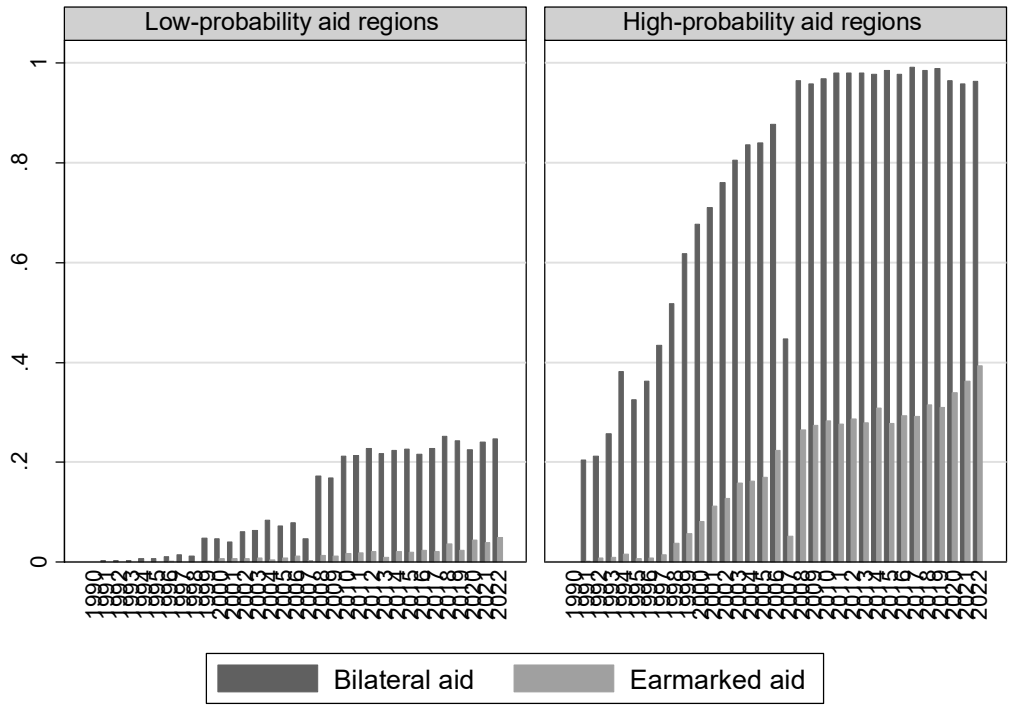

*Notes*: The figure shows no apparent trend differences in the outcome variable across exposure groups. Membership in different exposure groups is based on the median probability of receiving bilateral aid. The figure is qualitatively similar for the other subnational development outcomes. This supports the parallel trends assumption of the shift-share design.

**Figure A2**: Trends in cumulative number of regions assisted by aid




*Notes*: Dark gray bars show the (mean) cumulative number of regions assisted by a donor through bilateral aid, light gray bars through earmarked aid. The trends are similar across exposure groups; it is therefore not possible that unobserved characteristics between exposure groups can undermine instrument validity—even if the trends in the high-exposure regions are relatively similar to the trends of the endogenous treatments (Figure A3) and the outcome variables (Figure A1) as they are controlled for by constituent fixed effects in the first stage. This lends support to the notion that there are no omitted variables correlated with aid exposure that drive the results.

**Figure A3**: Trends in endogenous treatment variables



**Box A1:** Assessing the stringency of earmarked activities

As the geospatial analysis in this paper only includes projects with a precisely defined geospatial location, the coding of earmarking stringency (as adopted in the Earmarked Funding Dataset) needs to be adjusted. Because all activities are already earmarked at the country level, the key criterion for assessing earmarking stringency is now whether an activity supports a broad theme or sector, or whether it targets a narrowly-defined intervention.

In line with the coding in the Earmarked Funding Database, I follow three steps to assess the earmarking stringency of the aid activities.

- The default assumption is that all earmarked activities are strictly earmarked, unless otherwise specified. This is reasonable because the overwhelming share of earmarked activities are indeed strictly earmarked (UN-MPTFO 2022).

- Second, I identify softly earmarked activities using information in the Earmarked Funding Dataset on whether the donor supports a named entity (such as any trust multi-donor trust fund or funding facility), or whether the project title signals support for a broader activity.

  Support for a named entity: Component 1 of the Earmarked Funding Dataset enlists 730 institutional sub-accounts established under a parent organization. Any support to these accounts is considered a softly earmarked contribution as donors are subject to the same rules as any other donors under the common administrative agreement. A key example is if a donor supports the Energy Sector Management Assessment Program (ESMAP) in Indonesia, a multi-donor trust fund program of the World Bank (*tf=1* and *sdtf=0* in the dataset would require coding a softly earmarked activity --> *soft=1*)

  Support for a broad theme: Component 2 includes the titles (*t*) and detailed project descriptions (*l*) of all activities. A softly earmarked activity would match the following preset keywords: *soft=1*: strmatch(lower(t), "*program*")+strmatch(lower(t), "*core *")+strmatch(lower(t), "*unearmarked*")+strmatch(lower(t), "*membership*")+strmatch(lower(t), "*thematic *")+strmatch(lower(l), "*program*")+strmatch(lower(l), "*core *")+strmatch(lower(l), "*unearmarked*")+strmatch(lower(l), "*membership*")+strmatch(lower(l), "*thematic *")>0

- A soft coding can be overrun if the project description entails any of the trigger keywords for strict earmarking*: soft=0:* strmatch(lower(t), "*project*")+strmatch(lower(t), "* earmarked *")+strmatch(lower(t), "* restricted*")+strmatch(lower(t), "*specifi* contr*")+strmatch(lower(l), "*project*")+strmatch(lower(l), "* earmarked *")+strmatch(lower(l), "* restricted*")+strmatch(lower(l), "*specifi* contr*")>0

*Source*: Reinsberg, B., Heinzel, M., & Siauwijaya, C. (2024). Tracking earmarked funding in international organizations: Introducing the Earmarked Funding Dataset. *Review of International Organizations* (doi: 10.1007/s11558-024-09548-1).