Check for updates

Received: 21 March 2023 | Revised: 21 August 2023 | Accepted: 22 August 2023

DOI: 10.1111/1758-5899.13270

RESEARCH ARTICLE

Does earmarked funding affect the performance of international organisations?

Bernhard Reinsberg[1,2] | **Christian Siauwijaya**[1]

[1]School of Social and Political Sciences, University of Glasgow, Glasgow, UK

[2]Centre for Business Research, University of Cambridge, Cambridge, UK

Correspondence
Bernhard Reinsberg, School of Social and Political Sciences, University of Glasgow, Glasgow UK.
Email: bernhard.reinsberg@glasgow.ac.uk

Funding information
UK Research and Innovation, Grant/Award Number: MR/V022148/1

Abstract

What determines the performance of international organisations (IOs)? We argue that funding structures remain an overlooked driver of IO performance. Over the past decades, donor governments have provided an increasing share of their contributions to IOs in the form of earmarked resources, which provide them with the opportunity to restrict the use of funding to specific themes, sectors, regions, countries or projects. This development has raised concerns about the ability of IOs to perform their duties, given that earmarked funds distort programme priorities away from recipient need and increase transaction costs for IO staff. Beyond qualitative evidence from case studies, concerns about the performance-related effects of earmarking remain untested in large-N analysis. Drawing on 64 performance assessments of 32 IOs from 2009 to 2021, undertaken by the Multilateral Organisation Performance Assessment Network, we employ latent factor analysis to extract measures of *process performance*—the extent to which IOs have rules, procedures and routines in place to plan strategically, manage operations efficiently, liaise with partners effectively, monitor results and promote institutional learning—and *outcome performance*—the extent to which IOs achieve results that are relevant, efficient and sustainable. Using multivariate regressions, we find that earmarked funding is negatively related to process performance. Our result is robust to alternative model specifications and an instrumental variable design that helps mitigate concerns about endogeneity. These results have important implications for our understanding of IO performance and policy implications for donors of IOs.

1 | EARMARKING AND THE PERFORMANCE OF INTERNATIONAL ORGANISATIONS

Considering perennial global development challenges, such as climate change, conflict, food crises, and pandemics, there is a need for international organisations (IOs) that can deliver on their mandates. Despite recognition that IOs matter for achieving the Sustainable Development Goals (SDGs), we know little about what determines their organisational performance. Existing research has focussed on the institutional rules, relational capabilities and contextual factors that IOs can leverage to fend off attempts of political influence from member states (Honig, 2019; Johnson, 2014; Lall, 2017). For example, Lall (2017) finds that de facto autonomy, as reflected in the expert knowledge and the alliances that organisations form with non-state actors, is positively related to IO performance. Honig (2019) finds that greater autonomy for field-level agents, specifically in difficult development contexts such as fragile states, is linked to greater performance of development projects. Most recently, Heinzel (2022) finds that staff experience systematically affects project performance. However, whether and how funding structures affect performance has not been systematically studied.

In this article, we examine how funding structures affect the performance of IOs with a mandate to

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.
© 2023 The Authors. *Global Policy* published by Durham University and John Wiley & Sons Ltd.

promote global development. Taking a principal-agent perspective, we assume that different funding structures afford donor principals with different possibilities to exert control. On the one hand, core contributions provide IOs with much flexibility to allocate funding as they see fit to accomplish their mandates. Under core funding contracts, donors must pool their contributions and exercise control rights collectively, which dilutes their individual ability to influence IO decisions. On the other hand, earmarked contributions allow donors to restrict the use of delegated funds to specific themes, sectors, regions, countries or projects (Graham, 2017b; OECD, 2011; Reinsberg, 2017). Earmarked funding contracts not only specify funding purposes and execution modalities but also require organisations to prepare reporting tailored to donor needs. Earmarked contributions therefore come with additional transaction costs and operational restrictions that may limit the performance of implementing IOs.

As earmarked funding has skyrocketed in recent years, worries have grown that IOs will lose their ability to effectively respond to global development challenges. The United Nations (UN) has repeatedly warned against the adverse effects of earmarking in its Quadrennial Comprehensive Policy Reviews. These adverse effects include the distortion of programme priorities, loss of legitimacy and insufficient cost recovery of earmarked programmes (UN, 2017). Qualitative evidence from case studies undergirds these concerns, showing for select IOs that earmarked funding distorts spending priorities, increases administrative costs and reinforces conflicts between organisational entities (Baumann et al., 2019; Reinsberg, 2023; Schmid et al., 2021). Following the initiative of the UN Secretary-General, donor governments committed (in the so-called Funding Compact) to reduce the share of (strictly) earmarked funding if UN agencies could demonstrate reforms towards financial transparency, system-level cooperation and a clearer focus on results. Yet, UN agencies are complaining that donors do not honour their commitments, despite agencies' demonstrably successful reform efforts. Donors do not seem to believe that more unearmarked funding would boost the performance of IOs.

We provide the first large-N analysis of the relationship between earmarked funding and organisational performance at the level of IOs. To measure performance, we draw on a unique dataset assembled from 64 assessments of 32 IOs from 2009 to 2021, undertaken by the Multilateral Organisation Performance Assessment Network (MOPAN). Using factor analysis, we reproduce the conceptual distinction in the MOPAN assessment framework between *process performance* and *outcome performance*. Using latent performance scores and data on organisational funding structures from various sources, we find that earmarked funding is negatively related to process performance but does not affect outcome performance. A one-standard deviation increase in earmarked funding is related to a decrease in process performance by over one third of its standard deviation. This result holds when controlling for alternative explanations, like different levels of IO autonomy affecting both earmarking and performance. Moreover, our results also hold under an instrumental variable approach that exploits sector-level *ex-ante* similarity of IOs to predict their exposure to earmarking, albeit at a lower level of statistical significance.

Besides addressing an important policy challenge, our article offers several contributions for scholarship on IOs. First, our work consolidates, corroborates and extends previous evidence drawn from case studies and literature on earmarked funding. We show that the adverse effect of earmarking holds across different types of agencies and when considering alternative explanations for variation in IO performance. Second, regarding literature on IO performance, we suggest a new measure of performance based on systematic coding of MOPAN evaluation reports. These reports have been underused in related research due to methodological challenges, specifically the evolving nature of the assessment approach and hence the lack of over-time comparability of assessments. We overcome this limitation by taking a data-driven approach that generates valid, reliable and reproducible measures of performance and that vindicates the conceptual

Policy Implications

- Policymakers should carefully consider the implications of funding structures for the performance of international organisations.
- Performance assessments should distinguish between 'outcome performance' (Does an organisation achieve relevant and sustainable results at scale?) and 'process performance' (Does an organisation have the rules, procedures and routines in place that enable it to deliver results?).
- Earmarked funding allows organisations to expand their operations but supports activities that are less relevant and less sustainable. To enable international organisations to deliver relevant and sustainable results, donors should ensure adequate financial resourcing of these organisations.
- Earmarked funding undermines the process performance of international organisations. To enhance the efficiency of international organisations in delivering on their mandates, donors should unearmark their contributions to these organisations.

17588899, 2024, 1, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/1758-5899.13270 by University Of Glasgow, Wiley Online Library on [26/02/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

distinction between outcome performance and process performance in the IO literature.

2 | LITERATURE REVIEW AND THEORETICAL EXPECTATIONS

2.1 | IO performance and its determinants

Interest in understanding the drivers of IO performance has grown in recent years, among scholars (Gutner & Thompson, 2010; Lall, 2017; Lundgren et al., 2023) and in policy circles (Australian Government, 2012; DFID, 2011; MOPAN, 2019b). Different definitions emphasise different aspects of performance, ranging from outputs, outcomes, to impacts (Gutner & Thompson, 2010; Karlas & Parízek, 2019; Sommerer et al., 2022; Tallberg et al., 2016). Outputs are the tangible products of IO decision-making, such as resolutions, policies and project documents. Outputs are necessary for achieving outcomes, such as policy changes by national governments and development projects implemented by aid agencies. Impact is measured using the indicators encapsulated in the SDGs, such as life expectancy, maternal health, environmental protection and good governance.

Most definitions of performance recognise that it is both a process and an outcome (Gutner & Thompson, 2010; Lall, 2017; MOPAN, 2019b). For example, Lall (2017, 245) defines the performance of IOs as 'the extent to which they achieve their stated objectives and do so in a manner that is cost-effective and responsive to a wide range of (public and private) stakeholders'. Following this distinction, we define *outcome performance* as the extent to which IOs achieve results that are relevant to ultimate beneficiaries, advance cross-cutting issues, and adhere to principles of efficiency and long-term sustainability. We argue that outcome performance requires large scale and high-quality outputs. IOs that produce irrelevant results at scale may be as low-performing as IOs with high-quality projects that only reach few beneficiaries. In turn, we define *process performance* as the extent to which IOs have rules, procedures and systems in place to plan strategically, manage operations efficiently, liaise with partners effectively and monitor results and institutionalise organisational learning (MOPAN, 2019b). If IOs have robust operational processes, rules and routines in place, they should be more likely to generate results, although this association may be imperfect.

Existing research has focussed on institutional rules, donor practices and organisational behaviours to explain variation in IO performance. IR scholars tend to focus on donor preferences. While some argue that donors must impart controls on agencies to prevent agency slack (da Conceição-Heldt, 2017), others contend that donors often seek influence over bureaucratic decision-making to advance their selfish interests, to the detriment of IO mandate achievement (Gutner, 2005; Honig, 2019; Lall, 2017; Lundgren et al., 2018). Hence, IOs may benefit from greater autonomy with respect to their potential to deliver results, provided that their preferences are aligned accordingly (Eckhard & Ege, 2016; Ege, 2020; Honig, 2019). Organisational theorists and public administration scholars emphasise IO-specific factors, including institutional structures, organisational culture and individual staff (Bayerlein et al., 2020; Bove et al., 2020; Eckhard & Ege, 2016; Hall & Woods, 2018; Heinzel, 2022). For example, some argue that executive heads matter for IO outcomes (Copelovitch & Rickard, 2021; Hall & Woods, 2018; Parizek & Stephen, 2021), while others focus on the motivations, incentives and experience of operational staff and 'street-level bureaucrats' (Eckhard & Parizek, 2022; Heinzel, 2022; Lipsky, 2010). Yet, others contend that policy concentration—the ability of IO staff to focus their attention on a few strategic policies—affects IO performance (Steinebach et al., 2022).

Funding structures are somewhat neglected as a potential driver of IO performance. Scholars have long recognised that insufficient resources can undermine IO performance (Brown, 2010), and a rich historical literature has documented budget crises of the UN and its predecessor (Hirschmann, 2021; Hüfner, 2017; Patz & Goetz, 2019). However, beyond the notion that resourcing matters, systematic knowledge on how funding structures influence organisational performance is lacking.

2.2 | Funding structures and organisational performance

Principal-agent models provide a useful point of departure for generating theoretical expectations on the relationship between earmarked funding and organisational performance. While extant literature emphasises how donors delegate aid programmes to international bureaucracies to benefit from their global capacity, technical expertise and policy credibility (da Conceição-Heldt, 2017; Fleischer & Reiners, 2021; Hawkins et al., 2006; Nielson & Tierney, 2003), delegation from donors can take different forms, with different implications for agent autonomy. While core funding is provided by member states as a 'collective principal', earmarked funding establishes multiple principals that contract the agent under varying terms. Therefore, earmarked funding increases the opportunities for donors to exercise control over agency activities.

While enhanced donor control may not necessarily undermine IO performance, a rich body of qualitative work argues that it does (Baumann et al., 2019; Reinsberg, 2016; Schmid et al., 2021; Staeger, 2023).

17588899, 2024, 1, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/1758-5899.13270 by University Of Glasgow, Wiley Online Library on [26/02/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

However, these studies only ever look at specific types of IOs. Some cover individual UN agencies (Reinsberg, 2023; Schmid et al., 2021) and the UN system as a whole (Baumann et al., 2019), others capture the World Bank (Reinsberg, 2016), while others study regional organisations (Engel & Mattheis, 2020). Moreover, previous studies focus on distinct earmarked funding instruments such as multi-donor trust funds in fragile states, rather than earmarked funding as a whole (Barakat, 2009; Barakat et al., 2012; Muchadenyika, 2016). Finally, studies scrutinised different outcomes, such as the adequacy of the funding mechanism, the relevance for country needs and overall efficiency. A clear theoretical grounding in IO literature is missing. For all these reasons, the extent to which results generalise remains unclear. Synthesising insights from previous studies, we identify several mechanisms that underpin the expected negative relationship between earmarked funding and organisational performance.

First, earmarked funding can affect outcome performance by distorting activities towards low-quality interventions. Because earmarked funds can be provided by a single donor, they can make IOs focus on donor darlings that promise private benefits to donors but neglect investments into the collective assets of the multilateral system. For example, Reinsberg et al. (2017) show that small trust funds tend to support middle-income countries, which are more commercially attractive to donors than the poorest countries. The flipside of donor-driven earmarked aid may be that beneficiaries benefit less.

Earmarked funding may also concentrate on sectors with more easily measurable results. In the development domain, this is evident in the fact that earmarked funding supports short-term projects, compared to the long-term projects under core funding. IO staff may thus prioritise interventions with poorer fit with local circumstances and limited sustainability (Eckhard & Parizek, 2022; Gerard, 2022; Masaki et al., 2021). In the humanitarian domain, earmarked funding has been shown to delay responses to emergencies because agencies do not have flexible funds to jump-start an emergency response. Conversely, where donors provide flexible funding, humanitarian responses are more effective. During the COVID-19 pandemic, IOs with flexible multi-year funding were able to pivot some of their activities towards health care and hygiene promotion, while further allowing them to pre-finance new programmes and pre-position stocks and assets (Metcalfe-Hough et al., 2021, 100). In peace-building, research has concurred that earmarked funding interventions are too short to generate sustainable impacts, do not rely on local knowledge and do not provide sufficient resources for policy learning (Campbell, 2018). Hence, by altering the substance of IO interventions in the above ways, earmarked funding lowers the quality of IO outputs. This quality loss may undermine outcome performance unless earmarked funding helps agencies mobilise significantly more funding that could scale up their operations.

Second, earmarked funding is more costly for IOs to manage, given the need for additional reporting and increased fundraising. These additional demands likely strain core processes, such as strategic planning, human resources, relationship building and knowledge management, thereby undermining process performance. In fact, a key driver of such pressures are the additional reporting demands by donors. By allowing donors to directly contract IO agencies, earmarked funding comes with additional donor requests for tailored reporting. For instance, the WHO alone prepared 2301 interim reports and 699 final reports to individual donors in 2016 (Achamkulangare & Tarasov, 2017, 30). Reporting is particularly challenging at the programme level, where IO managers of earmarked contracts face the challenge of combining various pieces of information from operational teams that may not be commensurable (Achamkulangare & Tarasov, 2017, 5). Getting donors the information they want may necessitate additional staff capacities, which would become unavailable for other organisational functions.

Donors can shop around for other implementing agencies, fostering a competitive environment in which IOs are pressured to invest into donor relations for securing earmarked resources—oftentimes by relying on their core resources. However, time invested in fundraising implies time taken away from substantive project activities (Graham, 2017a; Reinsberg, 2016; United Nations, 2019). In an environment of scarce funding but multiple potential IO implementers, each IO will strive to secure as much financial resources as possible, even if it compromises organisational processes. For example, a MOPAN review of UN-WOMEN explained, "[t]he preponderance of earmarked funds also results in the high use of short-term appointments which poses risks for staff continuity and capacity, especially at country level. Staff are highly committed to their work, but the level of workload poses a risk of burnout." (MOPAN, 2019a, 49).

Competitive pressures also incentivise relevant IO actors to withhold information from each other, as IO actors could gain a competitive advantage by using their private information to attract donor funding. This could lead to the fragmentation of development efforts across and within IOs. In fact, these dynamics are not limited to inter-agency relations, but also apply to intra-agency relations (Baumann, 2021; Reinsberg, 2016; Schmid et al., 2021). For example, in the case of UN-WOMEN, "the co-ordination function however is negatively affected by lack of resourcing" (MOPAN, 2019a, 74). In sum, if an agency receives a greater share of its budget as earmarked contributions, its process performance will be lower, compared to a less-earmarked organisation.

17588899, 2024, 1, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/1758-5899.13270 by University Of Glasgow, Wiley Online Library on [26/02/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

3 | DATA AND METHODS

3.1 | Organisational performance

We draw on the MOPAN to measure organisational performance. MOPAN is a multilateral institution with 22 member states, supported by a permanent secretariat hosted by the OECD (MOPAN, 2021a). As a joint multilateral initiative, MOPAN promises to generate relatively unbiased evaluations, contrary to bilateral assessments, which often explicitly consider the contribution of IOs to specific national development objectives. In addition, MOPAN has more extensive coverage of IOs than bilateral donor assessments and, importantly, benefits from a unified assessment methodology that generates comparable performance scores across different IOs. Evaluations draw on document reviews (meta-analysis of IOs' evaluation reports); online surveys (donors, beneficiaries and peer organisations); interviews (both headquarters and country-level); and consultations (IOs' headquarters) (MOPAN, 2019b, 5). This diversity of sources helps validate and triangulate results, thus limiting bias in assessments while facilitating the use of standardised assessment indicators across IO with heterogenous features, histories and other contextual specificities.

Under its current methodology, the MOPAN evaluation grid includes 12 key performance indicators (KPIs) in five areas, each underpinned by several micro-indicators (MIs). The first four areas—strategic management, operational management, relationship management and performance management—gauge the extent to which IOs have rules, procedures and systems in place to devise aid strategies, manage operations efficiently, liaise with partners effectively and monitor results. These four areas seek to capture procedural aspects of organisational performance. The fifth area gauges the effectiveness of IOs in terms of their actual achievement of results, including relevance to country needs, efficiency of delivery, contribution to cross-cutting goals and sustainability (Table 1).

We encode detailed performance scores up to the MI level from the technical appendices of all available evaluation reports on the MOPAN website. KPI scores are computed as the average over constitutive MI scores and can fall within the following bands: highly satisfactory (3.01–4), satisfactory (2.01–3), unsatisfactory (1.01–2) and highly unsatisfactory (0.00–1). Thus far, MOPAN has completed 64 assessments of major development IOs from 2009 to 2021. Although some IOs were evaluated several times, MOPAN advises against comparing scores from different assessment cycles. Indeed, over-time comparisons of agencies are not possible because the MOPAN assessment grid evolved over time in two ways. First, the make-up of assessment areas changed. Under the so-called Common Approach, agencies were rated across 26 KPIs, based on over 70 MIs. Under the so-called MOPAN 3.0 approach, agencies have been rated across 12 KPIs, underpinned by over 60 MIs. Second, MOPAN re-designed the scoring scale in the transition from the Common Approach to the current MOPAN 3.0 approach, which makes it impossible to compare ratings across assessment cycles. Intermittently, the MOPAN 3.0* approach kept the new scoring grid, while the MOPAN 3.1 approach adopted a more demanding scoring within the existing scale of the MOPAN 3.0 approach. This was done because evaluators wanted to keep potential for differentiation at the top even though agencies had improved performance across the board. Fortunately, MOPAN makes untransformed scores available side by side the more stringent ones. In the appendix, we provide a mapping of the assessment criteria (Table S1) and assessment scales (Table S2) across the two major MOPAN approaches.

Despite these over-time changes, however, both MOPAN assessment methodologies—the Common Approach and the subsequent approaches based on MOPAN 3.0—seek to assess organisational performance with respect to strategic management, operational management, relationship management and knowledge management, as well as results attainment. We can therefore assume that all assessments capture the concept of organisational performance to some extent. Hence, the continuity in the overall understanding of performance allows us to construct a latent measure of performance. Given how MOPAN assessments have evolved, we do not attempt to compare agencies across different assessment cycles. Instead, we compute a continuous rank that places an organisation relative to its peers assessed in the same cycle. For example, the UN Office for the Coordination of Humanitarian Affairs (UN-OCHA) ranked almost two standard deviations below the average process performance score in its 2015 assessment (PP = –1.87) but improved by over one standard deviation (PP = –0.75) in its 2020–21 assessment.

In the following, we describe how we computed these continuous scores. Importantly, because the scoring grids are incommensurate, we perform factor analysis separately for the two main assessment cycle families. In a first step, we assess the internal consistency of all KPIs by testing whether the MIs underlying a given KPI load onto a single factor. We confirm this is indeed the case, suggesting that the MOPAN assessment framework is internally consistent. Second, we tried to run confirmatory factor analysis on all MIs, ignoring the information about their grouping into pre-defined KPIs. This was not possible because the number of MIs exceeds the number of observations, causing an insufficient degrees-of-freedom problem.[1] For the MOPAN 3.0 assessments, we therefore used the 12 KPIs as inputs to the factor analysis. We obtained these KPIs by taking the simple average across the associated MIs.

17588899, 2024, 1, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/1758-5899.13270 by University Of Glasgow, Wiley Online Library on [26/02/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

TABLE 1 MOPAN performance assessment framework.

Top-level concept	Performance area	Key performance indicators (KPIs)	Micro-indicators (MIs)
Process performance	Strategic management	KPI 1: Organisational and financial framework	MI 1.1: Long-term vision MI 1.2: Organisational architecture MI 1.3: Support to normative frameworks MI 1.4: Financial frameworks*
		KPI 2: Structures for cross-cutting issues	MI 2.1a: Gender equality MI 2.1b: Environment MI 2.1c: Governance MI 2.1d: Human rights MI 2.1e: *Cross-cutting theme relevant to the IO*
	Operational management	KPI 3: Relevance and agility	MI 3.1: Resources aligned to functions MI 3.2: Resource mobilisation* MI 3.3: Decentralised decision-making MI 3.4: Performance-based human resource management
		KPI 4: Cost effective and transparent systems	MI 4.1: Decision-making MI 4.2: Disbursement MI 4.3: Results-based budgeting MI 4.4: International audit standards MI 4.5: Control mechanisms MI 4.6: Anti-fraud procedures
	Relationship management	KPI 5: Relevance and agility in partnership	MI 5.1 Alignment MI 5.2 Context analysis MI 5.3 Capacity analysis MI 5.4 Risk management MI 5.5 Design includes cross-cutting issues MI 5.6 Design includes sustainability MI 5.7 Implementation speed
		KPI 6: Partnership and resources	MI 6.1: Agility MI 6.2: Comparative advantage MI 6.3: Country systems MI 6.4: Synergies MI 6.5: Partner coordination MI 6.6: Information sharing MI 6.7: Accountability MI 6.8: Joint assessments MI 6.9: Knowledge deployment
	Knowledge management	KPI 7: Results focus	MI 7.1: Results-based management focus MI 7.2: Results-based management applied in strategies MI 7.3: Evidence-based targets MI 7.4: Effective monitoring systems MI 7.5: Performance data applied
		KPI 8: Evidence-based planning	MI 8.1: Evaluation function MI 8.2: Evaluation coverage MI 8.3: Evaluation quality MI 8.4: Evidence-based design MI 8.5: Poor performance tracked MI 8.6: Follow-up systems MI 8.7: Uptake of lessons
Outcome performance	Results	KPI 9: Achievement of results	MI 9.1: Results deemed attained MI 9.2: Benefits for target groups MI 9.3: Policy or capacity impact MI 9.4: Gender equity results MI 9.5: Environment results MI 9.6: Governance results MI 9.7: Human rights results MI 9.8: *Results on IO-specific cross-cutting issue*

17585899, 2024, 1, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/1758-5899.13270 by University Of Glasgow, Wiley Online Library on [26/02/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

TABLE 1 (Continued)

Top-level concept	Performance area	Key performance indicators (KPIs)	Micro-indicators (MIs)
		KPI 10: Relevance to partners	MI 10.1: Target groups MI 10.2: National objectives MI 10.3: Coherence
		KPI 11: Results delivered efficiently	MI 11.1: Cost efficiency MI 11.2: Timeliness
		KPI 12: Sustainability of results	MI 12.1: Sustainable benefits MI 12.2: Sustainable capacity MI 12.3: Enabling environment

Note: Terms in *italics* are our own labels, although MOPAN uses them implicitly. MOPAN computes MIs as the simple averages of several items. KPIs are computed as simple averages over constitutive MIs. Asterisked MIs are those deemed potentially endogenous. We remove them in a robustness check.

TABLE 2 Factor analysis of micro-indicators with available data.

	Eigenvalue	Difference	Proportion	Cumulative
Factor1	4.100	1.672	0.546	0.546
Factor2	2.428	1.745	0.324	0.870
Factor3	0.682	0.256	0.091	0.961
Factor4	0.426	0.090	0.057	1.017

Note: The data only cover assessments from MOPAN 3.0 onwards. The constitutive variables in the factor analysis are KPIs. Output curtailed after four factors.

For the Common Approach assessments, using the 23 KPIs was not possible due to insufficient degrees of freedom. Therefore, we aggregated KPIs on process performance into assessment areas, taking simple averages of the KPIs associated with any given area. Third, we performed latent factor analyses with these variables for the Common Approach and for the later MOPAN approaches.

Table 2 shows the results of the confirmatory factor analysis for the MOPAN 3.0 methodologies. The Eigenvalue criterion suggests retaining two factors. Table 3 shows the loadings for these factors. The first factor loads positively on all KPIs, suggesting that all KPIs indeed measure the same latent concept that can be called 'organisational performance'. The second factor loads negatively on the KPIs of the four performance areas relating to procedural aspects, while loading positively on the four KPIs relating to results. This factor can therefore be called 'outcome performance'. It is remarkable that the factor analysis perfectly recovers the dimensions of the MOPAN assessment framework, which already foresees a distinction between 'process performance' and 'outcome performance'. In hindsight, this corroborates the validity of the MOPAN scores in measuring organisational performance. In the supplemental appendix, we show the results of the factor analysis for all evaluations between 2009 and 2014 under the Common Approach. We again obtain a two-factor solution (Table S3). All performance areas correlate with the first factor. Results-based performance areas positively load onto the second factor, and all but one process-based performance areas load negatively onto the second factor (Table S4).

Having performed factor analysis on the evaluations under both approaches, we can now combine the normalised ratings to obtain continuous relative rankings of all organisations. In essence, the retained factors summarise how an organisation scores in comparison with its peers rated in the same assessment cycle, respectively with respect to process performance and outcome performance.

Figure 1 shows the latent performance estimates for all IOs based on their available MOPAN evaluations from 2009 to 2021. We obtain meaningful differences in average performance between IOs, given that performance estimates are consistent across evaluation cycles, except for some IOs.[2] Moreover, process performance and outcome performance appear to be independent of each other.[3] Examples of high-performing IOs with respect to organisational processes include the Asian Development Bank (ADB), the Inter-American Development Bank (IDB) and the Montreal Protocol Fund Secretariat (MLF). These organisations score at least one standard deviation above the average rating of all other organisations assessed in the same cycle. In contrast, some of the lowest-performing IOs include the International Organization for Migration (IOM) and the UN Industrial Development Organization (UNIDO), with performance scores of almost two standard deviations below the field. With respect to outcome performance, low-performing IOs include the ADB, the International Fund for Agricultural Development (IFAD) and the Global Partnership for Education (GPE). The top-performing IOs are the IOM, MLF and UNAIDS.

17585899, 2024, 1, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/1758-5899.13270 by University Of Glasgow, Wiley Online Library on [26/02/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

3.2 | Funding structures

Our interest is in whether funding structures, particularly the division between core funding and earmarked funding, affect the performance of organisations. We combine several data sources to minimise loss of observations while maximising data accuracy.

For earmarked funding, we draw on an updated version of the multi-bi aid data, originally introduced for earmarked funding from all OECD/DAC donors to over 300 IOs in 1990–2012 (Eichenauer & Reinsberg, 2017; Reinsberg et al., 2023). This dataset is the best-available option for our purposes due to its wide coverage—including the earmarked activities of all OECD/DAC donors with over 340 IOs in 1990–2020. Other sources have less extensive coverage. For instance, the UN Data Cube only covers UN entities, and the official websites of non-UN organisations did not consistently report on earmarked funding flows.

For core funding, we combine two sets of data sources. First, we draw on the CRS-based DAC table on members use of the multilateral system, which records the core contributions by DAC member states to IOs in 2011–2020 (OECD, 2021). Where this dataset has missing observations, we draw on data from the UN Department of Economic and Social Affairs (UN-DESA) which keeps proprietary records on funding structures for all UN entities in the development domain. Combining these data sources allows us to minimise loss of observations due to listwise deletion.

An important decision pertains to how to treat the budgets of development banks. Here we take all ODA-eligible donor contributions as reported to the DAC, as opposed to just the contributions to concessional windows, like the International Development Association (IDA). In practice, this makes little difference: For core funding, for example, the World Bank Group received $8.6 billion, while IDA received $6.8 billion in 2020. More importantly, because earmarked funding is technically recorded for the IBRD branch, it would be misleading to only use IDA funding for both funding structures. For the regional development banks, we take a similar approach and combine OECD/DAC donor contributions to both the lending arms, concessional windows and private-sector windows.

Figure 2 shows the funding composition in the budgets of all IOs based on the year prior to their most

TABLE 3 Factor loadings of two retained factors.

Variable	Factor1	Factor2	Uniqueness
KPI 1	0.809	−0.147	0.324
KPI 2	0.140	−0.503	0.728
KPI 3	0.832	−0.131	0.291
KPI 4	0.760	−0.236	0.367
KPI 5	0.665	−0.170	0.529
KPI 6	0.515	−0.271	0.661
KPI 7	0.726	−0.041	0.471
KPI 8	0.417	−0.675	0.371
KPI 9	0.329	0.687	0.421
KPI 10	0.439	0.649	0.387
KPI 11	0.461	0.549	0.486
KPI 12	0.484	0.573	0.438

Note: The data only cover assessments from MOPAN 3.0 onwards.

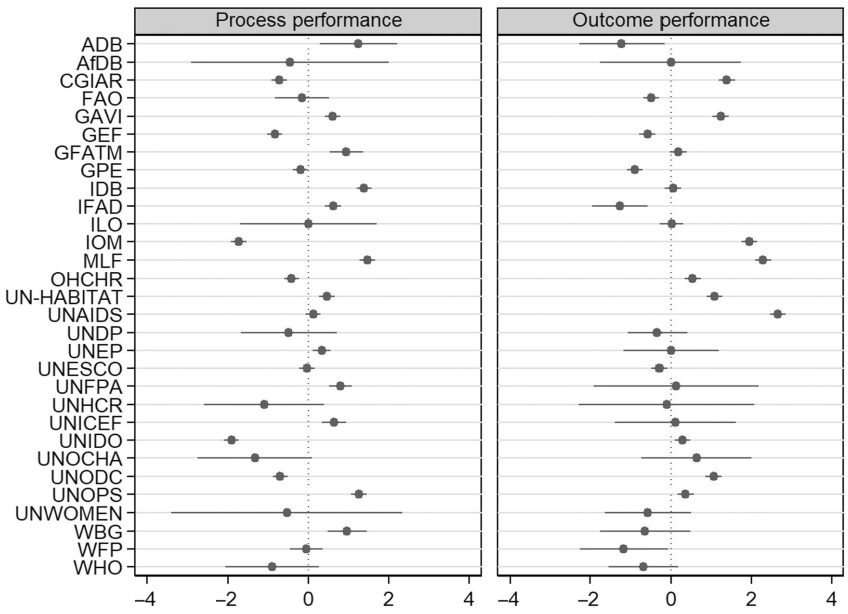


FIGURE 1 Organisational performance estimated from MOPAN evaluations. *Note*: The zero line delineates average performance. Filled dots show how the average evaluation of an organisation deviates from this average. Lines show the standard deviation across evaluation scores.

17585899, 2024, 1, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/1758-5899.13270 by University Of Glasgow, Wiley Online Library on [26/02/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

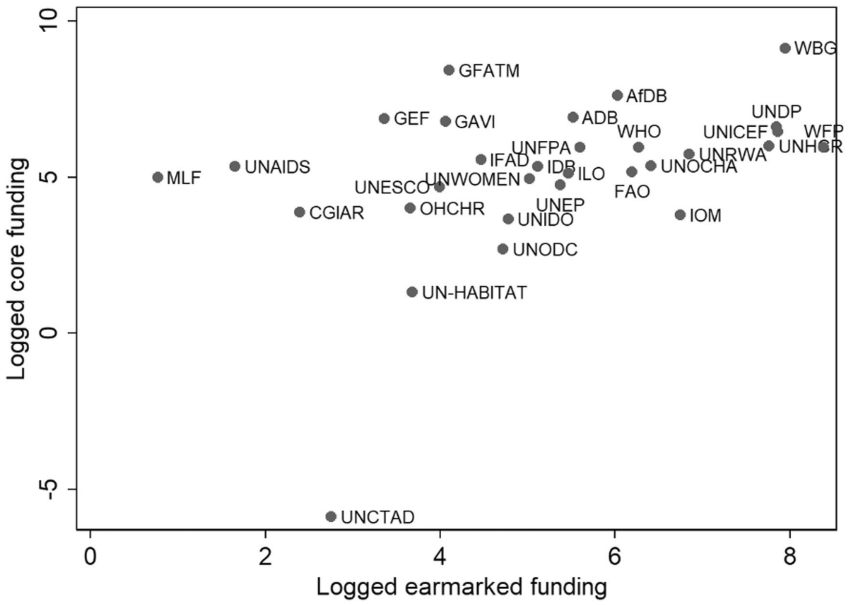


FIGURE 2 Funding structures across organisations. *Note*: The figure uses core funding and earmarked funding in the year prior to the most recent evaluation.

recent MOPAN assessment. The horizontal axis shows logged earmarked funding amounts, while the vertical axis shows the logged core funding amounts. The figure shows that the single largest organisation is the World Bank Group, with a relatively large share of core funding. In contrast, several UN organisations, like the UN Development Programme (UNDP), the UN High Commissioner for Refugees (UNHCR), The UN Children's Fund (UNICEF) and the World Food Programme (WFP), absorb similar levels of earmarked funding than the World Bank, but at far lower levels of core funding. At the other end of the spectrum, organisations like the MLF, UNAIDS and CGIAR have limited earmarked funds and relatively large pockets of core funding.

3.3 | Control variables

While the small number of observations poses limits to our inferences, we nonetheless conduct multivariate regressions which allow us to control for the most plausible confounders of the relationship between funding structures and organisational performance. We identify three sets of potential confounders.

First, we include a set of technical controls, given the way in which we obtained our performance data. Because ratings cannot be compared across MOPAN assessment cycles, we include a dummy variable for the more recent MOPAN methodologies. Essentially, this variable serves as a fixed effect that ensures that we only compare organisations within the same assessment cycle.

Second, we include proxies for organisational autonomy, given that more autonomous organisations may be better able to fend off pressures for earmarking while also being better performers. One source of autonomy lies with the Executive Head of an organisation—invariably referred to as Chief Executive Officer, Director-General or President (Copelovitch & Rickard, 2021; Hall & Woods, 2018; Parizek & Stephen, 2021; Saz-Carranza, 2015). We posit that organisations are more autonomous vis-à-vis member states if their executive heads have longer tenures, and if heads are appointed by a relatively large group of stakeholders. Executive heads appointed by the entire membership of an organisation can exploit preference heterogeneity of the donors and thereby carve out greater space for autonomous action. Another source of autonomy is the size of the Executive Board, the organ that manages the day-to-day operations of an organisation (Brown, 2010; Federo & Saz-Carranza, 2020; Haftel & Thompson, 2006). By the same logic, a larger board will find it more difficult to control the executive leadership and agree on sanctioning deviant behaviour.

Third, operational criteria, policy focus, organisational experience and geographical location may matter. The size of an organisation in terms of its total budget, its total staff count and its share of staff deployed in the field may be relevant. Larger organisations may be better able to exploit economies of scale and be able to afford fundraising departments. More staff under a given budget may matter, too, especially because many earmarked funding agreements come about locally between field staff and donor embassies (Baumann et al., 2019). We collect these pieces of information from publicly available websites and IO reports and take the logarithm to account for diminishing marginal utility. Donors may also perceive a greater

17588899, 2024, 1, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/1758-5899.13270 by University Of Glasgow, Wiley Online Library on [26/02/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

need for earmarking if an organisation is active in many policy areas. To measure policy scope, we count the (logged) number of CRS sector codes (OECD, 2021). To capture organisational experience, we compute the logged age of an organisation, based on IO website information. Finally, we collect information on headquarter locations, constructing a dummy for whether an organisation has a headquarter location in the Global South. Accounting for location may be important to account for varying accountability pressures on IOs given the presence of civil society networks (Dörfler & Heinzel, 2023). The appendix shows descriptive statistics of all variables (Table S5).

3.4 | Method

Our unit of analysis is the IO evaluation. Given that our outcome is continuous, we estimate linear regression using ordinary least squares. In robustness checks, we also use two-stage least squares regressions. We compute robust standard errors, which is an appropriate choice if, for example, assessment uncertainty has reduced over time.

4 | RESULTS

4.1 | Illustrative evidence

We proceed with bivariate plots to examine the relationship between earmarked funding and the two types of organisational performance. We expect that earmarking may affect performance only with a delay and earmarking shares may be subject to short-run fluctuations. Hence, we average earmarking shares over the 2 years strictly prior to the year of the evaluation.

Figure 3 presents the results. We find a significantly negative relationship between the average earmarked funding share and process performance. Some organisations with high earmarking and low performance include IOM, UNHCR and UNIDO. Some organisations with low earmarking and high performance include most vertical funds and the multilateral development banks. An example of an organisation that seems to deviate from this pattern is UNICEF, which performs better than expected, given its high earmarking share. Conversely, UN-OCHA—while being on the line of best fit with its 2020–21 evaluation—performed worse than expected, given its earmarking share in the 2015 evaluation. A closer look at the institutional history shows why: In 2015, UN-OCHA was found to be a procedurally poor performer, given 'weak systems for corporate risk management'—a finding that corresponded with an internal review that led to an organisational restructuring (MOPAN, 2021b, 12).

In the appendix, we present the findings of additional analysis suggesting that the negative relationship between earmarked funding and process performance is not driven by specific types of IOs, given that there is a similar negative relationship for both UN entities and non-UN entities (Figure S2). Figure 4 displays a weakly negative relationship between earmarked funding and outcome performance that is not statistically significant.

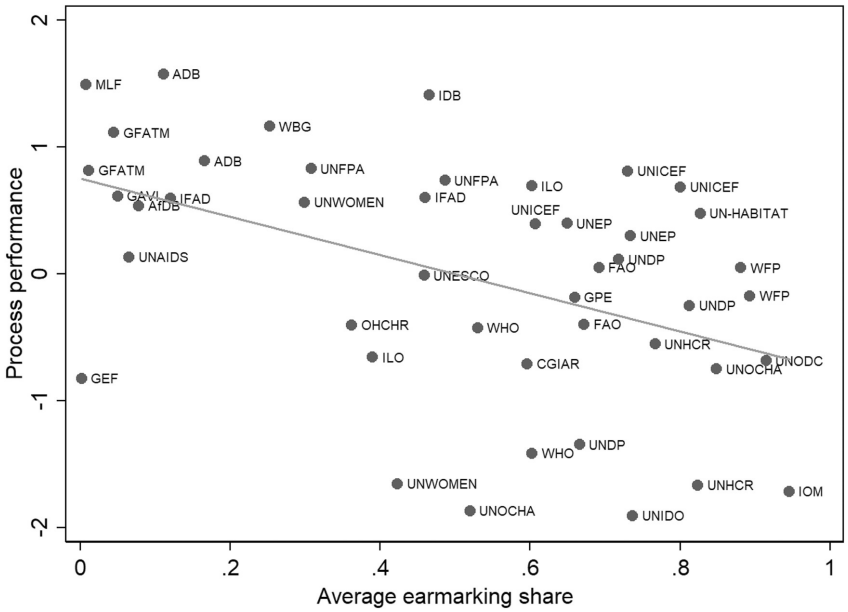


FIGURE 3 Earmarked funding and organisational performance. *Note*: Each filled dot represents one evaluation. The vertical axis shows process performance. The horizontal axis shows the share of earmarked funding in the total budget of an organisation averaged over the 2 years strictly before the evaluation year.

17588899, 2024, 1, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/1758-5899.13270 by University Of Glasgow, Wiley Online Library on [26/02/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

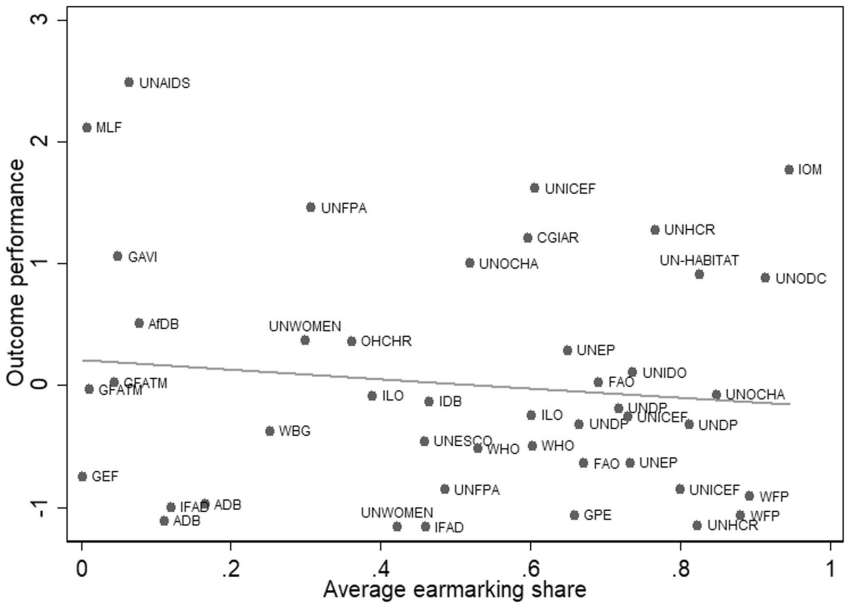


FIGURE 4 Earmarked funding and outcome performance. *Note*: Each filled dot represents one evaluation. The vertical axis shows outcome performance. The horizontal axis shows the share of earmarked funding in the IO budget averaged over the 2 years strictly before the evaluation year.

Taken together, the descriptive results suggest a negative correlation between earmarking and process performance, which can be interpreted in different ways. First, it could be that earmarked funding indeed undermines process performance, as documented by case studies (Baumann, 2021; Reinsberg, 2023; Schmid et al., 2021). Second, it could also be that underperforming organisations are more liable to earmarking, for example, because donors use earmarked funding to push for reforms. This is somewhat less plausible considering that organisations need flexible funding to support long-term reforms. Donors may also prefer to disengage entirely from organisations that they perceive as underperformers. For example, the United States withheld funding from IDA and withdrew from UNESCO (Hüfner, 2017; Patz & Goetz, 2019; Weaver, 2007). Australia left UNIDO in 1988, only to rejoin in 1992 when UNIDO could demonstrate that reforms promising greater performance were under way (von Borzyskowski & Vabulas, 2022, 22). Third, the relationship could still be spurious due to omitted variables. We will address this challenge through multivariate analysis.

4.2 | Multivariate analysis

Table 4 examines the relationship between earmarked funding and process performance for different sets of control variables. Our earmarking variable is the share of earmarked funding in the overall IO budget in the year before the evaluation. Across model specifications, we obtain a significantly negative relationship between the share of earmarked funding in an organisation and its process performance. In the barebones model with just assessment-cycle fixed effects, an increase in the share of earmarked funding by 30 percentage points (or one standard deviation) is related to a decrease in process performance by 0.397 (95% CI: 0.133–0.660) (or over one third of its standard deviation). The coefficient magnitudes are similar in the other models.

The following example helps to illustrate the substantive effect, provided the estimates can be causally interpreted. Consider the case of WFP, an organisation that is heavily dependent on earmarked funding. If WFP managed to reduce its earmarked funding share from 91.9% to 61.9%, its process performance would increase from −0.581 to −0.188. Of course, it may be implausible to argue that such drastic change in funding shares would be possible within the current budget envelope. The third model can answer this question as it includes the total IO budget as a control. This model suggests that to nullify the performance-enhancing effect of reduced earmarking in the WFP budget, the total budget would need to fall from its current $4.8 billion to approximately $134 million. The performance gains from reducing earmarking are likely to over-compensate any performance losses due to possible reductions in IO budgets, given that budget reductions would need to be unrealistically high.

Endogeneity is a concern. Performance itself may drive earmarked funding, or an omitted variable may drive both. To address this concern, we employ an instrumental variable design. We argue that for a given IO, the competitive pressure from organisations with

17588899, 2024, 1, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/1758-5899.13270 by University Of Glasgow, Wiley Online Library on [26/02/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

TABLE 4 Earmarked funding shares and process performance.

	(1)	(2)	(3)	(4)
Earmarked funding share	−1.322*** (0.435)	−1.017** (0.425)	−1.311** (0.534)	−1.006** (0.456)
MOPAN 3.0 onwards	−0.025 (0.309)	−0.043 (0.298)	0.045 (0.348)	0.013 (0.331)
Executive tenure over 5 years		−0.067 (0.267)		−0.339 (0.276)
Executive selection by assembly		−0.574* (0.302)		−0.755** (0.324)
Logged budget			0.106 (0.097)	0.099 (0.112)
Field share			−0.237 (0.456)	0.080 (0.489)
Logged number of sectors			−0.165 (0.247)	−0.326 (0.257)
Logged age			0.413 (0.372)	0.794* (0.428)
Headquarters in Global South			0.503 (0.570)	0.486 (0.553)
Observations	46	46	46	46
R-squared	0.171	0.232	0.248	0.359

Note: OLS regression with robust standard errors in parentheses. Outcome variable is process performance. Significance levels: *$p < 0.1$; **$p < 0.05$; ***$p < 0.01$.

TABLE 5 Earmarked funding shares and process performance using instrumental variables.

	(1)	(2)	(3)	(4)
Earmarked funding share	−1.501** (0.711)	−1.314* (0.697)	−1.534** (0.774)	−0.921 (0.621)
MOPAN 3.0 onwards	−0.024 (0.299)	−0.041 (0.282)	0.062 (0.322)	0.007 (0.301)
Executive tenure over 5 years		−0.074 (0.255)		−0.337 (0.247)
Executive selection by assembly		−0.497 (0.344)		−0.769** (0.307)
Logged budget			0.094 (0.102)	0.102 (0.100)
Field share			−0.221 (0.432)	0.079 (0.430)
Logged number of sectors			−0.104 (0.269)	−0.350 (0.265)
Logged age			0.418 (0.333)	0.796** (0.379)
Headquarters in Global South			0.453 (0.572)	0.503 (0.509)
Observations	46	46	46	46
R-squared	0.168	0.225	0.245	0.359
Kleibergen–Paap *F*-statistic	16.687	17.829	14.985	15.100

Note: IV/2SLS regression with robust standard errors in parentheses. Earmarked funding share is instrumented using the weighted mean of cosine similarities of sector spending profiles across all IOs in the sample in the year before the evaluation. Outcome variable is process performance. Significance levels: *$p < 0.1$; **$p < 0.05$; ***$p < 0.01$.

similar policy mandates is a good predictor of earmarked funding. If there are more competitors that—in the eyes of donors of earmarked funds—could perform the tasks at hand, then donors have leverage to force an organisation to accept earmarked funding. This argument establishes the relevance of the instrument. At the same time, this instrument is plausibly excludable, as competitive pressure will affect performance primarily through fundraising activity.[4] Instrument validity would be threatened if competition had a direct effect on performance, which is likely controlled by the count of sectors and other IO characteristics. To develop a continuous measure of competitive pressure, we construct sector profiles of multi-bi aid for each organisation and compute the (weighted) average cosine similarity over all organisation pairs. A sector profile consists of the budget shares across all CRS three-digit sectors (excluding non-allocable expenses). The cosine similarity is a relational measure that ranges from zero (for completely divergent sector profiles) to one (for identical sector profiles). The weighted average of these cosine similarities, where weights correspond to the size of organisations in terms of earmarked budgets prior to the evaluation year, gives the competitive pressure in the multilateral system facing a specific agency with a specific sector profile. Table 5 confirms that the instrument is highly correlated with earmarked funding shares ($F > 16$). We find that a higher share of earmarked funding seems to negatively affect process performance in the short term. Substantively, an increase in the earmarked funding share by 30% tends to reduce process performance by up to 0.46.

We now explore whether funding structures have an impact on outcome performance. Table S6 shows no relationship between earmarked funding and outcome

17585899, 2024, 1, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/1758-5899.13270 by University Of Glasgow, Wiley Online Library on [26/02/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

performance, for both OLS regression and instrumental variable regression. We could interpret these results in two ways. Substantively, the results could imply that worse process performance does not materialise in worse results. This could be because whether an organisation achieves results may be co-determined by extraneous factors, such as recipient-country behaviours. Methodologically, it might be difficult for evaluators to assess results across different organisations at a global level. To address this possibility, there is a need for complementary research at a more fine-grained level of analysis such as countries, regions or projects.

4.3 | Robustness tests

We probe robustness of our findings in the appendix. Specifically, we probe alternative explanations for our findings. One could argue that the type of organisation matters for the type of funding it receives. Similarly, larger IOs in terms of member states—especially when their preferences are poorly aligned—may have a greater share of earmarked resources. To address these possibilities, we control for whether the organisation has norm-making functions, as opposed to merely operational functions. Our main results are unaffected (Table S7). We measure the number of member states from the COW IGO dataset (Pevehouse et al., 2021) and IO websites. We measure the heterogeneity of member state preferences based on their UN General Assembly voting patterns (Bailey et al., 2017). Our main results are unaffected (Table S8).

In addition, we probe robustness to different measures and model specifications. Considering different lag structures, we find a strongly significant negative relationship between earmarked funding and process performance when using the average earmarking share in the 2 years prior to the evaluation year rather than the once-lagged earmarking share (Table S9).

We also probe alternative (and arguably cruder) measures of earmarking practice. Specifically, we count the (logged) number of earmarked projects and the (logged) number of donors in the three-year period before the evaluation. These measures tend to be negatively related to process performance with weak statistical significance if we do not control for IO characteristics (Table S10).

Furthermore, we probe a different variant of process performance that excludes two potentially endogenous MIs in the latent factor analysis. While these indicators are related to funding structures, they gauge the extent to which IOs mobilise, manage and report non-core resources in a way that supports the implementation of their mandate. Fortunately, our results do not hinge on the inclusion of these indicators (Table S11).

What is more, we use evaluation-year fixed effects instead of assessment-cycle fixed effects, which blocks arbitrary confounders in a given evaluation year. Our results are substantively similar and even more statistically significant for the instrumental variable analysis (Table S12). We also probe a simpler instrument: the (logged) number of sectors in which an organisation is active. IOs with broader portfolios may be more earmarked, given that donors may wish to support only some areas within these portfolios. At the same time, policy scope should not be directly related to performance. We obtain qualitatively similar results (Table S13).

Finally, given concerns with mixing data sources, we deploy a single-source measure of earmarked funding shares computed from UN data sources. This means that all non-UN entities drop from the sample. Even within the smaller sample, however, there is a significantly negative relationship between earmarked funding and process performance (Table S14).

5 | DISCUSSION AND CONCLUSION

This article examined how earmarked funding affects the performance of international development organisations. Earmarked funding allows donors to restrict their contributions to international organisations to specific themes, sectors, countries or projects, thereby limiting the autonomy of these organisations to spend their budget as they see fit.

To measure performance across organisations, we synthesised micro-level data from the MOPAN, which has conducted 64 evaluations of 32 IOs from 2009 to 2021. Due to changing assessment methodologies and assessment scales, the data have so far not been used for comparative evaluations of IO performance. To unlock the wealth of evidence enshrined in the MOPAN evaluations for comparative analysis, we conducted latent factor analysis, confirming the existence of two dimensions of performance. One measures *process performance*, understood as the extent to which IOs have rules, procedures and systems in place to plan strategically, manage operations efficiently, liaise with partners effectively and monitor results. The other is *outcome performance*, capturing the extent to which organisations achieve results that are relevant to ultimate beneficiaries, advance cross-cutting issues and adhere to principles of efficiency and long-term sustainability. While we extracted these dimensions empirically, they are reflected in relevant theoretical work on IO performance (Gutner & Thompson, 2010; Karlas & Parízek, 2019; Lall, 2017; Tallberg et al., 2016).

17588899, 2024, 1, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/1758-5899.13270 by University Of Glasgow, Wiley Online Library on [26/02/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Using latent performance scores and data on organisational funding structures from various sources, we found evidence that earmarked funding is negatively related to process performance but not outcome performance. This result was robust against different model specifications controlling for alternative explanations. Specifically, we confirmed that the results are not driven by differences in organisational autonomy, institutional features and contextual variables. Moreover, the relationship was not driven by specific types of organisations, such as the UN system.

We probed whether our estimates could be causally interpreted using an instrumental variable design. We argued that the competitive pressure facing an organisation determines its incentives to raise earmarked funds with donors. Where an organisation has a similar sectoral portfolio to its peers, it faces greater competitive pressure and therefore increases its fundraising. We found this instrument to be predictive of earmarked funding and recovered a significantly negative effect of earmarked funding on process performance.

Before discussing wider implications, we note the limitations of our study. With just 64 assessments, our tests might be underpowered, which could explain the null finding on outcome performance. Alternatively, the funding–performance link might be tenuous due to factors outside the control of IO staff. Moreover, measures of outcome performance at the organisational level might be too noisy to pick up systematic variation. Future research below the organisational level could address these concerns. Given our focus on the organisational level, we were limited to relatively sparsely controlled models that could only account for the most pertinent confounders. With regard to generalisability, we note that MOPAN covers the most important IOs, for which changes in funding structures could have tremendous performance effects. Although these IOs are not representative of the universe of IOs, their funding structures vary considerably, which increases our confidence that we would obtain similar results in a larger sample if data were available. There might also be concerns about our purely data-driven approach to measuring performance using factor analysis. Yet, this approach is appropriate because we could not map indicators between the Common Approach and the current one in a satisfactory manner, and we verified that the MOPAN Secretariat did not attempt such mapping either. Even if we could perfectly map indicators across vintages, we would still face the problem of how to convert scores between the Common Approach and the current one. Therefore, for the purposes of large-N analysis, we have used the rich information encoded in the MOPAN documents in the most pertinent way. That said, we believe there is potential for complementary qualitative comparative analysis of the MOPAN reports.

Our findings hold implications for related literature. They confirm the conclusions of recent research showing that member states hold the strings in the governance of contemporary IOs. However, different funding types are a subtler (and less studied) mechanism of influence compared to the withholding of contributions, populist contestation and membership withdrawal, which can prompt IO decline (Debre & Dijkstra, 2021; Eilstrup-Sangiovanni, 2021; Lake, 2021; von Borzyskowski & Vabulas, 2022). Future research could examine how earmarked funding affects development impacts, IO legitimacy and IO policy responsiveness. Future work could also examine how different types of earmarked funding affect different performance dimensions. Another area of future work pertains to the relationship between earmarked funding and organisational autonomy, which may require disaggregating international bureaucracies and considering different levels of analysis.

Our research also has important policy implications. Our findings reveal a paradox. They show that the same set of countries that push for more efficiency and better performance in the multilateral system simultaneously undermine it by earmarking their contributions. Earmarked funding—despite expanding IO resources—undermines the routines, processes and practices that enable organisations to deliver on their mandate. To increase process performance, an obvious solution would be to reduce the share of earmarking in IO budgets. In fact, over the past several years, the UN system has actively sought to nudge donors into providing more core resources while reducing earmarked funding and especially strictly earmarked funding. In April 2019, the UN and its donors agreed on the Funding Compact, in which UN agencies committed to embark on reforms aimed at increasing financial transparency, system-wide coherence and results focus, in exchange for more high-quality funding from donors (UN 2019). A similar agreement—the 'Grand Bargain'—was concluded between major donors and humanitarian agencies already in May 2016 (Metcalfe-Hough et al., 2021). According to many policymakers, donors did not live up to their promises, despite tangible UN reform progress (MOPAN, 2021a). This shows that unearmarking IO budgets faces many political obstacles. If not carefully managed, there may be a risk that donors lose domestic support for funding multilateral organisations altogether. Doubts about continuing past levels of engagement with multilaterals in recent aid papers of some donors, notably the United Kingdom, but also Norway—an erstwhile fervent supporter of the multilateral system—indicate that the next game for multilaterals may not be the fight for better funding, but the struggle for survival.

17585899, 2024, 1, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/1758-5899.13270 by University Of Glasgow, Wiley Online Library on [26/02/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

ACKNOWLEDGEMENTS

We thank participants at the BISA 2022 (Panel on 'International Institutions and Organizations: Complexity, Bureaucracy, and Funding') in Newcastle and three anonymous reviewers for helpful comments. This work was supported by a UKRI Future Leaders Fellowship (MR/V022148/1).

DATA AVAILABILITY STATEMENT

Replication data will be made publicly available on Harvard Dataverse (https://doi.org/10.7910/DVN/OVT16M).

ORCID

Bernhard Reinsberg https://orcid.org/0000-0001-7382-413X
Christian Siauwijaya https://orcid.org/0009-0005-6369-1907

ENDNOTES

[1] The outcomes of all preliminary validation steps are available on request.

[2] For example, process performance appears to vary significantly for ADB, ILO and UN-WOMEN.

[3] We confirm this with a bivariate plot in the appendix (Figure S2).

[4] The IV analysis would still be valid even if the instrument was not perfectly excludable. A Conley test indicated that the effect would still not vanish ($b \subset [-3.03, -0.01]$) even if competitive pressure had a direct effect on performance equivalent to about 30% of the estimated effect ($\gamma = 0.43$). This is a moderately high bar, given the model includes control variables and explains a moderately high share of the total variation.

REFERENCES

Achamkulangare, G. & Tarasov, G. (2017) *Review of donor reporting requirements across the United Nations system*. Geneva: UN Joint Inspection Unit.

Australian Government. (2012) *Australian multilateral assessment*. Available from: www.ausaid.gov.au/publications

Bailey, M.A., Strezhnev, A. & Voeten, E. (2017) Estimating dynamic state Preferences from United Nations voting data. *Journal of Conflict Resolution*, 61(2), 430–456.

Barakat, S. (2009) The failed promise of multi-donor trust funds: aid financing as an impediment to effective state-building in post-conflict contexts. *Policy Studies*, 30(2), 107–126.

Barakat, S., Rzeszut, K. & Martin, N. (2012) *What is the track record of multi-donor trust funds in improving aid effectiveness?*

Baumann, M.-O. (2021) How earmarking has become self-perpetuating in UN development co-operation. *Development Policy Review*, 39(3), 343–359.

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*. New York: UNDP, pp. 106–109.

Bayerlein, L., Knill, C. & Steinebach, Y. (2020) *A matter of style: organizational agency in global public policy*. Cambridge: Cambridge University Press.

Bove, V., Ruffa, C. & Ruggeri, A. (2020) *Composing peace: mission composition in UN peacekeeping*. Oxford: Oxford University Press.

Brown, R.L. (2010) Measuring delegation. *Review of International Organizations*, 5(2), 141–175.

Campbell, S. (2018) *Global governance and local peace: accountability and performance in international peacebuilding*. New York: Cambridge University Press.

Copelovitch, M.S. & Rickard, S.J. (2021) Partisan technocrats: how leaders matter in international organizations. *Global Studies Quarterly*, 1(1), 1–14.

da Conceição-Heldt, E. (2017) Regaining control of errant agents? Agency slack at the European Commission and the World Health Organization. *Cooperation and Conflict*, 52(4), 469–484.

Debre, M. & Dijkstra, H. (2021) Institutional design for a post-liberal order: why some international organizations live longer than others. *European Journal of International Relations*, 27(1), 311–339.

DFID. (2011) *Multilateral aid review: ensuring maximum value for money for UK aid through multilateral organisations*. Available from: http://www.nao.org.uk/wp-content/uploads/2012/09/1213594.pdf

Dörfler, T. & Heinzel, M. (2023) Greening global governance: INGO secretariats and environmental mainstreaming of IOs, 1950 to 2017. *Review of International Organizations*, 18(1), 117–143.

Eckhard, S. & Ege, J. (2016) International bureaucracies and their influence on policy-making: a review of empirical evidence. *Journal of European Public Policy*, 23(7), 960–978.

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. Available from: https://doi.org/10.1080/13876988.2020.1813032

Ege, J. (2020) What international bureaucrats (really) want: administrative preferences in international organization research. *Global Governance*, 26(4), 577–600.

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.

Eilstrup-Sangiovanni, M. (2021) What kills international organisations? When and why international organisations terminate. *European Journal of International Relations*, 27(1), 281–310.

Engel, U. & Mattheis, F. (2020) *The finances of regional organizations in the global south*. Abingdon: Routledge.

Federo, R. & Saz-Carranza, A. (2020) A typology of board design for highly effective monitoring in intergovernmental organizations under the United Nations system. *Regulation and Governance*, 14(2), 344–361.

Fleischer, J. & Reiners, N. (2021) Connecting international relations and public administration: toward a joint research agenda for the study of international bureaucracy. *International Studies Review*, 23(4), 1230–1247.

Gerard, K. (2022) "Doing aid chains differently." *Development policy review* (August 2021): 1–23.

Graham, E.R. (2017a) Follow the money: how trends in financing are changing governance at international organizations. *Global Policy*, 8(August), 15–25.

Graham, E.R. (2017b) The institutional design of funding rules at international organizations: explaining the transformation in financing the United Nations. *European Journal of International Relations*, 23(2), 365–390.

Gutner, T. (2005) Explaining the gaps between mandate and performance: agency theory and World Bank environmental reform. *Global Environmental Politics*, 5(2), 10–37.

Gutner, T. & Thompson, A. (2010) The politics of IO performance: a framework. *Review of International Organizations*, 5(3), 227–248.

Haftel, Y.Z. & Thompson, A. (2006) The Independence of international organizations: concept and applications. *Journal of Conflict Resolution*, 50(2), 253–275.

17585899, 2024, 1, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/1758-5899.13270 by University Of Glasgow, Wiley Online Library on [26/02/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Hall, N. & Woods, N. (2018) Theorizing the role of executive heads in international organizations. *European Journal of International Relations*, 24(4), 865–886.

Hawkins, D.G., Lake, D.A., Nielson, D.L. & Tierney, M.J. (2006) *Delegation under anarchy: states, international organizations, and principal-agent theory*. Cambridge: Cambridge University Press.

Heinzel, M. (2022) International bureaucrats and organizational performance: country-specific knowledge and sectoral knowledge in World Bank projects. *International Studies Quarterly*, 66(2), sqac013.

Hirschmann, G. (2021) International organizations' responses to member state contestation: from inertia to resilience. *International Affairs*, 97(6), 1963–1981.

Honig, D. (2019) When reporting undermines performance: the costs of politically constrained organizational autonomy in foreign aid implementation. *International Organization*, 73(1), 171–201.

Hüfner, K. (2017) The financial crisis of UNESCO after 2011: political reactions and organizational consequences. *Global Policy*, 8(5), 96–101.

Johnson, T. (2014) *Organizational progeny: why governments are losing control over the proliferating structures of global governance*. Oxford: Cambridge University Press.

Karlas, J. & Parízek, M. (2019) The process performance of the WTO trade policy review mechanism: peer-reviewing reconsidered. *Global Policy*, 10(3), 376–384.

Lake, D.A. (2021) The organizational ecology of global governance. *European Journal of International Relations*, 27(2), 345–368.

Lall, R. (2017) Beyond institutional design: explaining the performance of international organizations. *International Organization*, 71(2), 245–280.

Lipsky, M. (2010) *Street-level bureaucracy: dilemmas of the individual in public service*. New York: Russell Sage Foundation.

Lundgren, M., Squatrito, T. & Tallberg, J. (2018) Stability and change in international policy-making: a punctuated equilibrium approach. *Review of International Organizations*, 13(1), 547–572.

Lundgren, M., Tallberg, J., Sommerer, T. & Squatrito, T. (2023) When are international organizations responsive to policy problems? *International Studies Quarterly*, 67, sqad045.

Masaki, T., Parks, B.C., Faust, J., Leiderer, S. & DiLorenzo, M.D. (2021) Aid management, trust, and development policy influence: new evidence from a survey of public sector officials in low-income and middle-income countries. *Studies in Comparative International Development*, 56, 364–383.

Metcalfe-Hough, V., Fenton, W., Willitts-King, B. & Spencer, A. (2021) *The grand bargain at five years: an independent review*. London: ODI. Available from: https://cdn.odi.org/media/documents/GB_2021_WEB_YabmhpF.pdf

MOPAN. (2019a) *MOPAN 2017–18 assessments: UN women*. Paris: Organization for Cooperation and Development (OECD). Available from: https://www.mopanonline.org/assessments/unwomen2017-18/Final_assessment_UN_Women_2019_02_14.pdf

MOPAN. (2019b) *MOPAN 3.0 methodology manual*. Paris: Organization for Cooperation and Development (OECD). Available from: http://www.mopanonline.org/ourwork/ourapproachmopan30/Methodology-Manual-FINAL.pdf

MOPAN. (2021a) *Is this time different? UNDS Reform: progress, challenges and opportunities. lessons in multilateral effectiveness* (June).

MOPAN. (2021b) *United Nations Office for the coordination of humanitarian affairs (OCHA)*. Paris: Organization for Cooperation and Development (OECD).

Muchadenyika, D. (2016) Multi-donor trust funds and fragile states: assessing the aid effectiveness of the Zimbabwe multi-donor trust fund. *Journal of International Development*, 28, 1337–1357.

Nielson, D.L. & Tierney, M.J. (2003) Delegation to international organizations: agency theory and World Bank environmental reform. *International Organization*, 57(2), 241–276.

OECD. (2011) *2011 DAC report on multilateral aid*. Paris: DCD/DAC(2011)21/FINAL.

OECD. (2021) *Creditor Reporting System (CRS)*. stats.oecd.org (October 20, 2021)

Parizek, M. & Stephen, M.D. (2021) The long march through the institutions: emerging powers and the staffing of international organizations. *Cooperation and Conflict*, 56(2), 204–223.

Patz, R. & Goetz, K.H. (2019) *Managing money and discord in the UN: budgeting and bureaucracy*. Oxford: Oxford University Press.

Pevehouse, J.C., Nordstrom, T., McManus, R. & Jamison, A.S. (2021) Tracking organizations in the world: the correlates of war IGO version 3.0 datasets. *Journal of Peace Research*, 57(3), 492–503.

Reinsberg, B. (2016) The implications of multi-Bi financing on multilateral agencies: the example of the World Bank. In: Mahn, T., Negre, M. & Klingebiel, S. (Eds.) *The fragmentation of aid: concepts, measurements and implications for development cooperation*. Basingstoke: Palgrave Macmillan, pp. 185–198.

Reinsberg, B. (2017) Trust funds as a lever of influence at international development organizations. *Global Policy*, 8(5), 85–95.

Reinsberg, B. (2023) Earmarked funding and the performance of international organizations: Evidence from food and agricultural development agencies. *Global Studies Quarterly*. Available from: https://doi.org/10.1093/isagsq/ksad056

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

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.

Saz-Carranza, A. (2015) Agents as brokers: leadership in multilateral organizations. *Global Policy*, 6(3), 277–289.

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.

Sommerer, T., Squatrito, T., Tallberg, J. & Lundgren, M. (2022) Decision-making in international organizations: institutional design and performance. *Review of International Organizations*, 17(4), 815–845.

Staeger, U. (2023) Resource mobilization in security partnerships: explaining cooperation and coercion in the EU's partnership with the African union. *Cooperation and Conflict*. Available from: https://doi.org/10.1177/00108367221147785

Steinebach, Y., Knill, C., Kaplaner, C. & Bayerlein, L. (2022) Bureaucratic agency and policy performance in international organizations. In: Knill, C. & Steinebach, Y. (Eds.) *International public administrations in global public policy*. London: Routledge, pp. 125–145.

Tallberg, J., Sommerer, T., Squatrito, T. & Lundgren, M. (2016) The performance of international organizations: a policy output approach. *Journal of European Public Policy*, 23(7), 1077–1096.

United Nations. (2017) *Report of the secretary-implementation of general assembly resolution 67/226 on the quadrennial comprehensive policy review of operational activities for development of the United Nations system (QCPR): funding analysis*.

United Nations. (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: United Nations. Available from: https://documents-dds-ny.un.org/doc/UNDOC/GEN/N19/096/71/PDF/N1909671.pdf?OpenElement

17585899, 2024, 1, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/1758-5899.13270 by University Of Glasgow, Wiley Online Library on [26/02/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

von Borzyskowski, I. & Vabulas, F. (2022) When is withdrawal from international organizations deadly? In *APSA 2022 annual meeting, Montreal*.
Weaver, C. (2007) The World's Bank and the Bank's world. *Global Governance*, 13(4), 493–512. Available from: http://www.jstor.org/stable/27800679

AUTHOR BIOGRAPHIES

Bernhard Reinsberg is a Reader in Politics and International Relations at the University of Glasgow and a Research Fellow in Political Economy at the Centre for Business at the University of Cambridge. His research is on the political economy of development cooperation, specifically funding practices of international development organisations.

Christian Siauwijaya is an economics graduate and a Research Assistant in Politics and International Relations at the University of Glasgow. Christian is interested in organisational performance, economic development and political economics.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

Appendix S1

How to cite this article: Reinsberg, B. & Siauwijaya, C. (2024) Does earmarked funding affect the performance of international organisations? *Global Policy*, 15, 23–39. Available from: https://doi.org/10.1111/1758-5899.13270

17585899, 2024, 1, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/1758-5899.13270 by University Of Glasgow, Wiley Online Library on [26/02/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License