The Review of International Organizations
https://doi.org/10.1007/s11558-026-09639-1

Peer-to-peer accountability and knowledge production in OECD DAC peer reviews

Alice Iannantuoni[1] · **Simone Dietrich**[2] · **Bernhard Reinsberg**[3,4]

Received: 22 May 2025 / Accepted: 22 July 2026

Abstract
The OECD Development Assistance Committee (DAC) relies on peer reviews to promote accountability in development cooperation. For peer review to function as an accountability mechanism, it must first satisfy a core requirement: peer review evaluations must accurately reflect donor practices. We examine whether DAC peer reviews interpret donor behavior on tied aid in ways that are consistent with observable aid-tying practices and shared DAC standards. Using a newly collected corpus of 143 DAC peer reviews from 1994 to 2020, we find a strong, systematic relationship between the reviewed donor's aid-tying practices and the resulting content of the review documents. This linkage is driven primarily by explicit recommendations to untie aid and discussions of the standard, rather than by positive recognition of aid tying. Contrary to concerns that peer reviewers may intentionally soften evaluations and criticism to protect their own interests, we find no evidence that reviewers who themselves exhibit higher levels of aid tying produce evaluations that deviate from actual practice. Our findings demonstrate that DAC peer reviews generate credible peer evaluations that strengthen accountability dynamics in development cooperation. These insights contribute to literatures on peer reviews in global governance, the politics of international organizations, and foreign aid.

Keywords OECD DAC · International organization · Peer review · Development cooperation · Norms and standards · Information production · Accountability

JEL Classifications F35 · F53 · D73 · D83 · O1

Responsible Editors: Stephen Chaudoin, Christina Davis, and Christoph Mikulaschek

✉ Simone Dietrich
simone.dietrich@unige.ch

[1] Université Catholique de Lille, Lille, France

[2] University of Geneva, Geneva, Switzerland

[3] University of Glasgow, Glasgow, UK

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

Published online: 05 August 2026

1 Introduction

International organizations (IOs) collect, standardize, interpret, and disseminate information, often as a tool to promote accountability. One example is the peer review process in the OECD's Development Assistance Committee (DAC), where foreign aid donors periodically assess each other regarding their performance against IO standards for effective aid-giving. A key practice against which donors are assessed through peer review is aid tying, which requires recipients to procure goods and services from donor-country providers. This practice is widely seen as undermining aid effectiveness because it restricts recipient governments' ability to decide how to best employ aid funds to serve local priorities and often increases project costs (Annen & Kosempel, 2009; Ganga, 2025).

The OECD DAC has identified aid tying as a problem for development cooperation at least since 1965. Then, the OECD DAC acknowledged the "desirability of a common effort by donor countries to reduce progressively the scope of aid tying with a view ultimately to removing procurement restrictions to the maximum extent possible" (OECD, 1965). Although donors have made considerable progress in untying their aid since the 1960s, recent data and NGO reporting suggest that this practice is again on the rise (e.g., Eurodad, 2021). Figure 1 shows that aid tying among traditional donors steadily declined from the late 1980s until the mid-2000s. Newer donors who joined the DAC in 2010[1] have historically tied larger shares of their aid. Moreover, in recent years, even major traditional donors report increases in aid tying. For example, in 2022, only about half of Italy's aid was reported as untied, while top donors like the United States and Japan have increased tied aid to about one-third of their overall aid, respectively.

These developments raise questions about the DAC's ability to hold donors accountable. Some scholars attribute renewed aid tying to bureaucratic politics or a broader shift toward national donor interests (Bau & Dietrich, 2025, 2026; Carcelli, 2024; Dietrich, 2021). Others argue that it reflects weaknesses in the DAC's accountability mechanisms (Rabinowitz, 2024). Like many standard-setting international organizations, the DAC lacks strong formal enforcement tools to compel compliance with its standards. Instead, it relies primarily on bureaucratic instruments—most prominently peer review—to monitor donor behavior and promote accountability.

As accountability tool, a peer review assesses donor performance relative to shared standards and identifies areas where practices diverge from agreed norms and standards. By interpreting donor behavior and issuing recommendations, peer reviews generate signals around which political actors, including domestic aid actors, civil society, or other donors, can mobilize to exert policy pressure. In this sense, peer review operates as a bureaucratic tool through which the DAC translates observable donor practices into structured qualitative evaluations and recommendations. International organizations often exercise authority through such bureaucratic processes of interpretation and classification (Barnett & Finnemore, 2004), and similar dynam-

[1] These are the Czech Republic, Estonia, Hungary, Iceland, Korea, Lithuania, Poland, the Slovak Republic, and Slovenia.

Fig. 1 Average aid tying practice over time. *Notes:* The figures shows the percentage of bilateral aid commitments that are not marked as "untied" according to the OECD DAC ODA tying status data set (DAC7b). The solid black line represents the DAC donor average; the dotted line represents the average of the DAC top-five bilateral donors (France, Germany, Japan, United Kingdom, and United States); and the dashed line the average of new donors who joined the DAC since 2010. While these data are technically available since 1979, this donor-reported information is understood to be trustworthy starting in the 1990s

ics appear in other DAC reporting systems where bureaucratic classification rules shape how donor behavior is represented (Bau et al., 2026b).

For peer review to support accountability, however, peer examiners must accurately interpret donor practice. To that end, the peer review team relies on DAC guidelines that structure and standardize how reviewers are expected to evaluate donor behavior. Yet, peer review is inherently relational: how donor governments evaluate the behavior of others may be shaped by their own aid practices. For example, donors that engage in aid tying themselves may have incentives to evaluate peers more leniently. Such bias would undermine the accountability function of peer review, which depends on credible and politically independent evaluations. Accurate evaluation is therefore a prior condition for accountability to function. As consensus around the norm against aid tying weakens, disagreements over how the norm should be applied may further influence how reviewers interpret donor behavior.

Despite these concerns, evidence on the content of DAC peer reviews is limited. We address this gap by examining whether DAC peer review reports accurately interpret donor aid tying practice. We focus on DAC norms around tied aid because it represents one of the most politically sensitive and economically consequential issues in development cooperation. Donor governments have long faced strong domestic pressures to use foreign aid as a tool to promote national export interests (Ganga, 2025). Given these domestic pressures, tied aid constitutes a most-likely case for uncovering political influence in DAC peer reviews. If peer reviewers were inclined to shield fellow donors from criticism, this is an area where such bias should be most visible.[2] Specifically, we test whether the qualitative evaluations contained in peer reviews correspond to observable donor aid tying practices reported in OECD aid data. If peer review functions as intended, reviews should systematically reflect donor behavior. If instead evaluations are influenced by the practices of reviewing states, the relationship between donor behavior and peer review content should weaken.

A better understanding of how vulnerable DAC peer reviews may be to political influence from member states is important because it sheds light on the DAC's ability to address accountability deficits in development cooperation. When donor governments violate the shared aid untying standard, peer reviews should document this malpractice and make recommendations to reduce it (e.g., Carraro et al., 2019). If peer review content becomes politicized, however, the process risks losing its value as a bureaucratic mechanism for producing credible assessments of how donors do on aid tying. Such distortions would weaken peer-to-peer accountability and challenge the DAC's authority as a global standard setter.

We analyze the universe of 143 DAC peer reviews from 1994 to 2020 with a newly collected corpus of review documents, assessing the extent to which reviews discuss aid tying and whether or not they contain an explicit recommendation against it. We measure aid-tying practice with OECD data on Official Development Assistance (ODA) from DAC donors. We find support for the notion that the peer review process is indeed implemented following the guidelines set by the DAC review team: there is a positive and statistically significant relationship between a donor's actual aid-tying practice and whether its review contains more mentions of tied aid or a recommendation against aid tying. Distinguishing between praising, neutral, and negative mentions of the tied aid topic in peer reviews, we further show that this positive relationship is indeed stronger when we remove praising mentions from our analyses.

Our results contribute to several debates in the IO literature. First, we contribute to debates on the functioning and effectiveness of peer review (Carraro et al., 2019; Conzelmann, 2014; Groenendijk, 2009; Jongen, 2021; Terman & Voeten, 2018). Our analysis shows that the *content* of DAC peer reviews, with respect to the practice of aid tying, systematically reflects observable donor behavior, and is thus not distorted by political bias. In doing so, we complement previous work on the *assignment* mechanism of peer reviewers, which finds no evidence that examiners are assigned to review teams for political reasons across cases (Iannantuoni et al., 2025). Our findings therefore provide new evidence that peer review can generate credible information about state behavior.

More broadly, our findings contribute to our understanding of how IOs hold their member states accountable. Accurate information and interpretation of member state behavior are understood as preconditions for accountability. If DAC peer reviews were biased because of political interference, they would provide less accurate

[2] In turn, if we were to find little evidence of political influence in this domain we believe that other issues treated in the DAC peer review, including policy coherence for sustainable development, anti-corruption standards, or concessionality requirements, would be even less likely to be affected by bias insofar as they often involve more technical and less politically salient issues.

assessments and be less useful for monitoring and encouraging best practices in aid giving. Although peer review is a relational governance mechanism that potentially allows for peer influence, our results suggest that it can still function as a credible information mechanism. Unlike processes that create accountability through quantifiable indicators or grades (e.g., Honig & Weaver, 2019; Kelley & Simmons, 2021), DAC peer reviews rely on qualitative evaluations that highlight malpractice and issue recommendations for improvement. While this approach may be less confrontational than governance by numbers, it ensures that IO standards like aid untying remain in the spotlight and lead to debates about how development cooperation should evolve (e.g., Hale, 2017). Our measurement strategy captures the extent to which peer reviews respond to observed member state practice and could be applied to study accountability dynamics in other consensus-based regimes (e.g., the Paris Agreement).

2 Accountability in aid-giving through the DAC peer review

IOs are important actors in their own right that collect, categorize, interpret, and disseminate information (Barnett & Finnemore, 2004). They derive their power from their ability to define meanings and shape expectations—they "govern by information" (Von Bogdandy & Goldmann, 2008; Biermann & Siebenhüner, 2009; Jorgens et al., 2016; Busch & Liese, 2016). Through their role as information-brokers, IOs can facilitate cooperation among states by relaying (more) credible information about state preferences, state behavior, and the state of the world (Abbott & Snidal, 1998; Keohane, 1984; Koremenos et al., 2001). This informational authority also enables IOs to promote accountability among member states. In the words of Grant and Keohane (2005, 29), accountability "implies that some actors have the right to hold other actors to a set of standards, to judge whether they have fulfilled their responsibilities in light of those standards, and to impose sanctions if they determine that those responsibilities have not been met."

By providing information about and evaluation of what states do, IOs have an important role to play in helping stakeholders hold states accountable against their promises. IOs provide information to diverse audiences, including peers, domestic and foreign publics (Mikulaschek, 2023; Voeten, 2001). Information provided by IOs—including raw data, qualitative reports, and numerical rankings—can induce behavioral change in member states by facilitating market pressure, external discourse, and internal norms (Hale, 2008). For example, the literature on 'governance by numbers' emphasizes the powerful role of market pressure in promoting state compliance, given that numerical rankings make it easy for audiences to compare across states (Honig & Weaver, 2019; Kelley & Simmons, 2021; Morse, 2019; Simmons, 2000). Beyond numerical indicators and rankings, IOs also promote accountability through qualitative evaluation mechanisms such as peer review, where members of the peer review team interpret state behavior and assess compliance with shared standards.

For these accountability benefits to materialize, target audiences must be able to trust the information that is provided. While the rational-legal nature of IOs makes them generally trustworthy sources of information, the extent to which they can fulfill this role is contextually contingent. First, IOs are believed to provide more credible information where their secretariats enjoy greater autonomy and where their monitoring role is more comprehensive (Breen & Doak, 2021; Mitchell & Hensel, 2007; Sadeh & Porath, 2020). For example, Oksamytna (2018) shows how an autonomous UN peacekeeping staff devised a public information program even without a relevant mandate. The type of information that IOs collect also matters: The trend toward combining complex information into simple rankings not only gives extraordinary powers to IOs that produce them, but it also pushes ranked states to adopt policies of a questionable scientific basis via market pressure (Broome et al., 2018).

Second, IOs often depend on the willingness of member states to provide accurate information. However, certain types of states, specifically populist governments, do not share information or, if they do so, the information is biased or delayed. This behavior reflects the anti-elite sentiment of populists, who counteract the proper functioning of IOs by withholding information (Carnegie et al., 2024) or reducing the reporting on bureaucratic accountability tools set by the IO (Bau et al., 2026b). In addition, autocracies may be less likely to share information in a timely manner as they may fear opposition groups or elite factions exploiting such information to threaten their survival in office (Hollyer et al., 2014).

Third, scholars have also noted that geopolitical context matters (Reinsberg & Heinzel, 2026). The information transmission function of IOs generally works as intended when all states have pro-cooperation incentives. However, in a fractured world marked by geopolitical rivalry, information is more likely to be 'weaponized'—being used, manipulated, or withheld strategically to disadvantage geopolitical rivals (Farrell & Newman, 2019).

While the existing literature provides important insights into when and why information provision through IOs works as it should, our study draws attention to the design of the information-gathering process itself. When information collection relies primarily on intergovernmental mechanisms, states may have greater opportunities to exert influence over how information is produced and interpreted, potentially reducing its credibility. In contrast, secretariat-driven processes, especially those that grant secretariats independent powers to collect information, are more likely to yield high-quality information. Peer review occupies an intermediary position on this continuum. In peer review processes, states evaluate each other's performance against agreed norms and standards, while a secretariat facilitates the process and structures the resulting reports. These evaluations often combine hard information—encoded in verifiable data—with soft information derived from qualitative assessments that contextualize and compare state practices across cases (Conzelmann, 2014).

One prominent example of such peer review processes is the peer review mechanism of the OECD's Development Assistance Committee. The DAC is a consensus-based, standard-setting organization created to ensure that international development donors give aid in a recipient-friendly way, including the untying of aid, among other standards. At the same time, the DAC also seeks to enforce accountability by encouraging its members to comply with standards of aid giving so that they can structure donor-recipient interactions. However, like many other consensus-based IOs, the DAC has limited means available to hold its members accountable. It relies on a peer review system that asks members to evaluate each other's aid-giving practices against its development standards. While some have suggested that peer-to-peer accountability through peer review has worked to keep aid more recipient-friendly (Ashoff, 2013), others have argued on the basis of anecdotal evidence that recent increases in aid tying across DAC members may result, in part, from the DAC's failure to hold its members accountable through its peer review system (Rabinowitz, 2024). Knowing how vulnerable the DAC peer review system may be to political influence is important for further research that examines the nature of information generated through the process as well as, more broadly speaking, the dynamics of peer-to-peer accountability in international development.

DAC peer reviews are conducted by a core team of OECD officials together with representatives of two member states designated as peer examiners.[3] The Director and Deputy Director of the Development Co-operation Directorate (DCD) officially lead the DAC peer reviews, while the Head of the Reviews, Results, Evaluation, and Development Innovation Division (RREDI) guides and oversees the peer review process as a whole. Within RREDI, the Peer Review Team Leader manages the schedule of reviews and provides quality assurance for individual reviews. A lead analyst is assigned to each peer review with responsibility for managing the review, including the assessment of data collected through the country's aid statistics.

While member states broadly agree on how reviews are to be conducted, the OECD's Development Cooperation Directorate safeguards the development of a comprehensive peer review framework and its impartial, technocratic implementation against the DAC's standards and practices. Specifically, as regards the DAC standard on aid tying, peer evaluations assess members against the standard enshrined in the 2001 Recommendation on Untying Official Development Assistance.[4]

During the review process, peer examiner states and DAC representatives monitor the reviewed state's recent record on aid tying and report on it in the review document. For example, Finland's 2012 review discusses aid untying in a dedicated section titled, "Dealing with an Increasing Share of Tied Aid," and includes explicit references suggesting that Finland "has made progress against all but one of the Paris Declaration indicators: the exception being untied aid" (OECD DAC 2012). The 2017 South Korea review includes a roadmap for moving toward greater untying and a recommendation asking for better reporting of aid-tying status (OECD DAC 2017).[5] While OECD reporting systems already document donor aid practices, peer reviews perform an important interpretive function by evaluating those practices relative to shared DAC standards.

These examples suggest that the peer review content should be an accurate reflection of the donor's underlying practice, suggesting that the peer review follows the technocratic ideal, free of particularistic member state interests. Indeed,

[3] Examiners are represented by high-level officials with at least ten years of work in their aid agency across management, operational, or policy positions, who can share experience from their history with the reviewed donor (OECD DAC 2020, p. 5).

[4] In parallel, these same standards have featured prominently on the agenda of successive international agreements on effective development cooperation including in Paris (2005), Accra (2008), Busan (2011), and Nairobi (2016), among others.

[5] The Appendix includes further illustrations of mentions and recommendations.

related research confirms that the *assignment* of peer reviewers—determining who reviews, whom, when, and where—largely follows the DAC's bureaucratic protocol and is insulated from political pressures that may lead to favoritism in the selection of reviewers (Iannantuoni et al., 2025).

From a theoretical perspective, we can derive competing expectations as to whether and when the DAC peer review can fulfill its accountability function. In their role as examiners, DAC review teams are expected to interpret donor behavior and assess whether it aligns with shared international standards on the basis of data from OECD reporting systems. In an optimistic view, we may expect peer review to work as intended, given that the organization has detailed peer review guidance, supported by a capable secretariat that strives to apply the rules consistently across all donors. Examiner states are required to follow guidance on how to conduct the evaluation, including donor self-assessments, data collection, the selection of analytical assessment criteria, the analysis of information and data, as well as communication and reporting requirements. Underlying a process-oriented technocratic approach to peer review is the belief that, by codifying the evaluation process, the organization can compare states on their compliance with standards and offer useful feedback for improvement. Our in-depth research into the DAC reveals that the peer review regime has been carefully codified, drafted, and revised in the form of an institutional protocol that maximizes the quality and impartiality of its procedures and outcomes.

> **Hypothesis 1**: The more a donor engages in the practice of aid tying, the more it is likely that the norm against aid tying will be assessed and discussed in its peer reviews.

In a more critical account, peer review is often portrayed as an example of relational governance where states evaluate other states. While the DAC may strive to conduct process-oriented peer reviews, these evaluations do not take place in a vacuum but inside a consensus-based political organization where donors do not always see eye to eye. As with other peer reviews, states have incentives to evade criticism of their behavior that could harm their reputation and induce audience costs (Hai, 2025; Keesing, 1998; Tingley & Tomz, 2022). Likewise, a large literature in political economy has established how states influence policymaking in IOs when their core interests are at stake (Bau & Dietrich, 2025; Dreher et al., 2009; Kilby, 2011; Stone, 2011). States exert influence over IOs primarily through financial leverage, often referred to as "power of the purse" (McLean & Park, 2025; Reinsberg & Heinzel, 2026; Zhang & Chen, 2025). More specifically for IFIs, powerful nations utilize their agenda-setting authority and informal clout to prioritize favorable loan conditions for key strategic allies (Anyiam-Osigwe & Vreeland, 2026; Dreher et al., 2009; Lim & Vreeland, 2013). Beyond funding, states strategically place their nationals in high-level positions to subtly steer policy recommendations toward domestic interests (Gui, 2024; Novosad & Werker, 2019; Reinsberg & Heinzel, 2026; Zaccaria & Reinsberg, 2025).

Our qualitative research on the DAC peer review process suggests that there are opportunities for examiner states to influence the language used in the evaluation and the adoption of explicit recommendations. In fact, the protocol includes meetings where the review team and the donor under review can discuss and contest the language and recommendations in a review draft. This draft is then finalized in a joint editing session before being released to the public. Because the DAC is a consensus-based organization, this kind of transparency and communication are essential. At the same time, closed-door discussions may allow states to promote their interests and shape the content of the resulting review documents.

Related research from other peer review mechanisms suggests that donors often interfere with due process, especially where their salient interests are concerned. For example, in his assessment of the Trade Policy Review Mechanism (TPRM) at the World Trade Organization, Keesing (1998, 30) attests that "[secretariat] staff often finds itself under considerable pressure from the country under review to give it favorable marks, and less pressure from any other quarter to give out tougher ones." In the context of the United Nations Universal Periodic Review, Terman and Voeten (2018) find that examiner countries tend to be less critical of their allies' records than they are of non-allied peers. We are cautious about drawing direct analogies from these peer reviews because their institutional processes differ from the DAC's. However, these differences are not categorical, but represent a continuum of authority, specifically regarding their insulation from donor pressure (Jongen, 2018). Despite the relative insulation of DAC peer review from donor pressure, we may still expect attempts of undue influence at the margins, especially where reviewers have a poor track record on untying aid and may therefore want to avoid being too critical of their peers.

Some might argue that politics is unlikely to affect peer review because the DAC is a relatively homogeneous organization where membership correlates with decisions to undertake economic reforms and geopolitical alignment (Davis, 2023). While we agree that a shared understanding of the need to solve transnational problems like poverty and ineffective aid can facilitate governance in the DAC, its members nonetheless disagree over technical approaches—and the extent to which they untie aid limits their ability to have domestic stakeholders benefit from it. These disagreements are openly discussed and contested in DAC working groups, and eventually become manifest in cross-donor differences in aid-tying.

Acknowledging the existence of heterogeneous preferences in aid-giving practices among DAC members, we see how donors may have incentives to shape how they are reviewed and how they review others. A member serving as a peer examiner whose government has a poor track record of complying with DAC standards may prefer to limit the extent to which reviews highlight non-compliance with those same standards. This implies that examiner states can use peer evaluation to advance their political interests. A similar argument—described as "glasshouse effect"—has been made in the context of the WTO TPRM, which "[a]t times… inhibits participants from throwing stones" (Keesing, 1998, 11). In our empirical analyses, we interpret statistically significant associations between the aid-tying practice of peer examiners and the content of peer reviews as evidence that the peer review process might be driven by donor concerns over how they are represented.

There are other mechanisms beside the "glasshouse effect" for why some reviewers would prefer to sweep problematic aid-tying practices of their peers under the carpet. For example, reviewers with a poor record in untying aid may prefer to be silent about similar practice by peer donors to prevent the mobilization of domestic interests. This motivation yields observationally equivalent predictions as the glasshouse effect, which focuses on the international reputation among policymaking elites. In contrast, we would expect limited variation in our results if either the norm against aid tying is weakening or if the technical aspects of its implementation are viewed as ambiguous. Under these conditions, the past record of the peer reviewer on aid tying would cease to be a relevant factor. In light of our discussion, we formulate our second hypothesis:

> **Hypothesis 2**: Tied aid discussion and explicit recommendations against aid tying in peer reviews are less likely to occur when peer examiner states also have a poor record on untying aid.

In the remainder, we bolster some key assumptions underlying our theoretical argument. In particular, we outline how our argument is specific to the peer review context, and why peer review is a useful mechanism for various stakeholders to fulfill their accountability demands.

In contrast to "governance by numbers," peer reviews avoid explicit rankings and are first discussed and authorized by examiner and reviewed countries. In addition, the DAC and its members understand peer review as a critical learning and policy coordination tool insofar as the reviews provide extensive hands-on opportunities for members to learn from each other. As the DAC has grown from less than ten original donors in 1960 to 32 members in 2023, the heterogeneity among members has markedly increased, arguably making the learning dimension of the peer review system ever more important. New and emerging donors, in particular, are thought to benefit from knowledge and experiences of their peers when it comes to building up their own domestic aid systems (Bau et al., 2025). To maximize learning through peer review, peer-to-peer evaluations should accurately reflect and discuss the state of affairs of the donor under review.

Contrary to the view that DAC peer reviews are merely "cheap talk," our extensive qualitative research suggests that they are resource-intensive instruments that foster donor accountability through significant international and domestic signaling. The process requires a heavy investment of time and capital from the OECD Secretariat, the reviewers, and the reviewed donors. For the donor under review, the burden of collecting comprehensive data across complex bureaucracies—ranging from small, resource-constrained agencies to the (erstwhile) sprawling, multi-agency systems of the United States—is substantial. This year-long commitment involves deep diplomatic engagement, headquarter assessments, and field visits[6] to aid programs in developing countries, ensuring that examiners fully grasp the specific operating environment of their peers.

Beyond the administrative costs, these reviews serve a vital domestic function by generating data for a variety of stakeholders. Official launch events and press releases disseminate findings to legislatures, civil society, and the media, providing a unique framework for national actors to benchmark their performance against international

[6] For a more in-depth discussion on the implications of these field visits for knowledge about what works in international development see Dietrich et al. (2025).

standards. For example, legislative committees use these reports to inform aid policy and budget decisions, while NGOs leverage the data to pressure bureaucracies or highlight issues like aid-tying (e.g., Eurodad, 2024). Ultimately, the peer review process functions as a high-stakes reputational mechanism: positive evaluations bolster agency credibility and national pride, while critical reports create public and political pressure for reform.

In the landscape of IOs, the OECD DAC peer reviews are often referred to as the "gold standard" of soft law governance. Several institutional features help explain this reputation.[7] First, the OECD Secretariat possesses high technical but low political autonomy. It manages the peer review process and drafts the review report while the findings must eventually be negotiated into a politically acceptable consensus. Second, members show a high willingness to provide accurate data, largely because the DAC's credibility is built on its role as the world's "official record-keeper" for development finance. In this context, disputes over reporting do not arise from failing to disclose aid activities but rather over the interpretation of what counts as Official Development Assistance, including in-donor refugee costs or certain security-related expenditures. Third, the DAC is a relatively homogeneous club, reducing the risk of conflict over geopolitical issues (Davis, 2023). While the DAC was historically a "like-minded" club of Western donors, it has successfully integrated Eastern European states that have been socialized into DAC practices. The peer review mechanism has had to adapt to varying levels of institutional capacity and different historical motivations for aid. Finally, DAC peer reviews combine technocratic evaluation with staged politicization: following closed-door peer review meetings where officials respond to detailed scrutiny, published reports often mobilize broader audiences down the line, including domestic and international advocacy groups that use OECD assessments to pressure governments. For instance, if the DAC recommends that a donor government should make more effort in untying its aid, domestic advocacy groups may use the peer review as a tool to pressure the donor government after the peer review report gets released. These institutional features make the DAC a useful setting in which to evaluate whether peer review generates credible assessments rather than politically distorted ones.

3 Data and methods

We are interested in whether the evaluations in DAC donor peer reviews on tied aid track donor aid behavior rather than parochial interests of reviewing countries. We study these questions with linear models, regressing the degree to which a peer review

[7] We derive a similar conclusion from applying a prominent accountability framework: Grant and Keohane (2005) highlight four critical drivers of accountability: clearly understood rules and norms, recognition of authority and the right of the reviewer to judge, mechanisms to gather behavioral data, and penalties for deviant behavior. The DAC peer review fulfills most of these conditions: the 2001 Recommendation on Aid Tying established a strong norm, underpinned by robust reporting and data collection that allows evaluators to judge donor practices and requires reviewed states to justify their practices in front of an intergovernmental panel. Although there are no hard sanctions for non-compliance, reputational costs can be high when peers criticize a state or NGOs mobilize against aid tying.

discusses the issue of aid tying and recommends increased compliance with the organization's recommendation to untie bilateral commitments onto (i) the reviewed donor's aid-tying track record and (ii) the aid-tying track record of the two peer examiners.

The key outcomes in our analyses are mentions of aid tying and the presence of an explicit recommendation against it in 143 reviews from 1994 to 2020. The consistency of aid-tying data across donors beginning in the early 1990s motivates our focus on this period (see Appendix for further descriptions on the universe of DAC peer reviews).

3.1 Recommendations on aid tying

We operationalize our primary outcome—the degree to which a peer review document addresses aid tying— using four distinct measures. These measures are informed by the internal structure of DAC documents and the political process through which they are produced.

Each peer review consists of two primary components. The first is the Secretariat's Report, a comprehensive assessment (approximately 5,000–30,000 words) co-produced by DAC officials and examiner country delegates. It provides technical depth, including data on the reviewed donor's aid scope and recipient country profiles. The second is the summary section, which is a concise opening (550–3,500 words) that highlights key findings and lists explicit recommendations for the donor.

The distinction between these sections is analytically significant due to the "peer review meeting" that concludes the process. Before official publication, the reviewed donor meets with DAC officials and peer examiners to discuss the draft. During this stage, the donor may challenge specific findings or negotiate edits.

We contend that the summary section carries higher political stakes than the Secretariat's Report insofar as its brevity and explicit recommendations ensure greater visibility and readership. Consequently, donors are likely to push back more vigorously against a formal recommendation in the summary than against similar critiques buried within the technical report. Peer examiners with poor compliance records may prefer that sensitive issues be relegated to the 'thick' of the Secretariat's Report rather than highlighted in the summary. If political negotiation influences the final text, these dynamics are most likely to manifest in the summary's language and recommendations.

To account for these structural and political nuances, we employ four specific outcome measures that respectively capture the prevalence of aid-tying mentions in (1) the full peer review document, (2) the Secretariat's Report, and (3) the summary section; and finally (4) whether or not the front-end summary includes explicit recommendations against aid tying.[8]

To compute the prevalence of discussion on aid tying, we utilize a dictionary of keywords[9] and extract all sentences centered around keywords from the peer review

[8] Research assistants manually identified whether each recommendation pertains to aid tying.

[9] `Tie," ``ties," ``untie," ``unties," ``*tied," and ``*tying." Qualitative checks make us confident that our approach does not leave much room for false negatives, and research assistants qualitatively assessed each tied aid mention to remove any false positives. See the Appendix for a thorough discussion.

documents. A qualitative reading of the sentences shows that the practice of aid tying is typically discussed to assess the reviewed donor's progress toward the untying goals set by the DAC's 2001 Recommendation. We additionally code for whether each mention contains praise for good or improved performance.

Of the 143 peer reviews in our sample, only two have no meaningful[10] tied aid mention: the 1997 review of the Netherlands, and the 2001 review of Belgium. The rest have between one and 69 tied aid mentions, and about 17 on average. In the appendix, we examine the distribution of aid tying recommendations over time: while the DAC made its first aid tying recommendation (to the UK) in 1997, all reviews included such recommendations already by 2000; the 2001 Recommendation on Tied Aid further consolidated this trend (Figure A2).

To examine where recommendations are placed in the review documents, we measure the prevalence of tied aid keywords over total word counts at review-document level. Figure 2 depicts the prevalence of tied aid mentions and how those mentions are distributed between the Secretariat’s Report and the summary section. On the horizontal axis, we list each of the 143 documents, ordered chronologically first and alphabetically second. We do not observe an obvious time trend; rather, there is considerable variation in the positioning of tied aid mentions across reviews.

3.2 Aid tying practice

The main explanatory variable is the reviewed donor’s actual track record on aid tying in the years leading up to the review. If DAC peer reviews are to serve as effective accountability tools, they must explicitly measure a donor’s adherence to DAC standards. We measure aid-tying practice with data made available by the OECD in Table DAC7b: Aid (ODA) Tying Status: we first construct a donor-year measure as the proportion of total bilateral commitments[11] explicitly marked as “untied,” and then subtract it from one, to obtain the donor-year proportion of tied aid.[12] We take this donor-year measure of aid-tying practice and construct its two-year average, for the year of a review and the year before it.[13]

We expect the reviewed donor’s track record to matter in explaining the variation in the content of review documents. However, we also consider whether peer examiners influence review outcomes. We use a measure of peer examiner aid-tying practice that mirrors that of reviewed donor aid-tying practice. We take the average of

[10] A mention that is not simply part of a glossary of definitions, but is genuinely part of the review-specific language.

[11] According to Table DAC7b documentation, this total does not include administrative costs and technical cooperation expenditure because these types of aid are exempt from the 2001 Recommendation.

[12] We follow advice from OECD CRS officials and calculate this measure starting from the amount of aid commitments explicitly labeled as “untied.” Everything that is not reported as untied is therefore considered tied: *Tied Aid* = *1* − *Untied Aid*. Note that it would also be possible to focus on amounts marked as “tied” or as “partially tied,” but these data are less complete. Figure 1 depicts the distribution of this variable.

[13] In the Appendix, we present robustness checks operationalizing the reviewed-donor practice as the percentage of tied aid in the year before the peer review and as a three-year average. We use a two-year average for our main specifications because it best resonates with our understanding of the timeline of the review process and how data are brought into the peer evaluation.

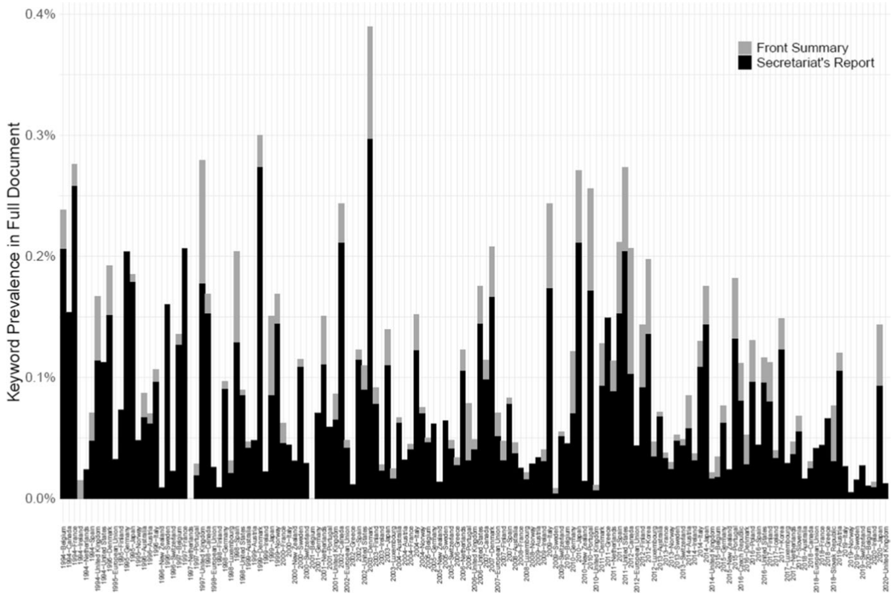

Fig. 2 Prevalence of tied aid mentions in full peer review documents. *Notes*: Review-level percentage of tied aid keywords over total word count–-showing considerable variation across reviews. Most tied aid mentions appear in the Secretariat's Report (in black) rather than in the shorter front-end summaries (in gray)

that measure (*Tied Aid* $_{t-1:t}$) across the two examiners, if data are available for both; if data are missing for either of the two examiners, we use the aid-tying practice of the other alone.

To further probe the plausibility of our mechanism, we test a secondary empirical implication of our theoretical framework. If indeed the DAC peer review system fairly assesses donors' compliance with its standards and best practices, the relationship between donor practice and review content should be different when it comes to content that praises compliance. As discussed, most mentions of aid tying in peer reviews occur in the context of either reporting the reviewed donor's compliance with the 2001 Recommendation or explicitly discussing a worsening in donor compliance. Sometimes, however, mentions occur in the context of praising the reviewed donors for improved compliance—i.e., praise for a reduction in aid tying.[14]

In the case of review content praising improved compliance, we expect the opposite relationship with donor practice: the more a donor ties its aid, the less likely it will be to receive praise for its compliance with the 2001 Recommendation. Conversely, if we were to remove these praising mentions from our sample, the positive relationship between aid tying practice and tied aid mentions should remain strong or become stronger.

[14] Research assistants manually identified whether each tied aid mention praised the reviewed donor for improved performance.

Table 1 Effect of reviewed donor practice on tied aid evaluation

	Tied aid mentions			
	Full document	Report	Summary	Recommendation
	(1)	(2)	(3)	(4)
Tied Aid $_{t-1:t}$	0.127***	0.117***	0.181*	0.562**
	(0.031)	(0.028)	(0.079)	(0.206)
Mentions in DCR $_{t-1}$	13.030	9.308	22.766	90.683
	(10.100)	(9.214)	(26.032)	(64.106)
Constant	0.055***	0.054***	0.086*	0.146
	(0.015)	(0.014)	(0.038)	(0.095)
Observations	116	116	116	105
R^2	0.145	0.143	0.052	0.086
Adjusted R^2	0.130	0.128	0.035	0.068

† $p<0.1$; * $p<0.05$; ** $p<0.01$; *** $p<0.001$

4 Results

Is the reviewed donor's aid-tying practice systematically associated with the prevalence of tied aid mentions in its subsequent peer review, as well as the presence of a specific recommendation against aid tying? Across four linear regression models presented in the top panel of Table 1, we find that the percentage of a donor's actual tied aid commitments in the years immediately before a review is a positive and significant predictor of the prevalence of tied aid mentions in that review. This effect exists when assessing the prevalence of tied aid mentions in the peer review document as a whole (Model 1), in its Secretariat's Report section (Model 2), and in its front-end summary section (Model 3), as well as when assessing the presence of a tied aid recommendation (Model 4). The effect size is particularly large for the recommendation outcome. For this outcome, an increase in the percentage of tied aid by a standard deviation is associated with an increase in the likelihood of a recommendation on tied aid by 13.4 percentage points—equivalent to over one-fourth of a standard deviation ($p<0.05$).[15] To put this result into context: a one-standard deviation increase in tied aid mentions in the OECD DAC's Development Cooperation Reports (DCRs) is only associated with an increase of 6.2% in the likelihood of a tied aid mention in the peer review document, and this relationship is not statistically significant.

Note that these results remain significant whether or not we include a control for the extent to which the DCRs discuss the topic of tied aid.[16] DCRs are the organization's annual flagship reports in which, for over 50 years, the DAC has aggregated information on trends and best practices in international development cooperation efforts. The fact that our findings are robust to the inclusion of this control indicates that we are not conflating general time trends related to the discussion of aid tying in the DAC donor community broadly with the donor-specific assessments that take place in the peer reviews.

[15] 1 SD for tied aid is 22.4%, which leads to a predicted change in probability of 13.4% which—given 1 SD of the outcome is 47.6%—translates into 28.2% of its standard deviation.

[16] We replicate our results without the DCR variable in the appendix (Table A5).

In the appendix, we probe the robustness of our main result for alternative measures of tied aid practice. Our main analysis uses a two-year average ($t-1$: t) of tied aid because it best reflects the timeline of the review process: the review team works on assessing the reviewed donor's record for a few months leading up to the review meeting, likely having access to data from those years. We demonstrate that the results are similar when using a three-year average ($t-2$: t) (Table A1), or the year before the peer review meeting ($t-1$) (Table A2).

For another alternative measure of the tied aid share, we use the flows explicitly reported as "tied aid" in DAC Table 7b. While all coefficients are positive, only those for the full document and the body of the report are statistically significant (Table A3). We obtain a similar result when combining the shares of "tied aid" and "partially untied aid" (Table A4).

To further test our mechanism, we examine whether the link between actual aid-tying and written mentions disappears when the peer review is specifically praising a donor. In Table 2, we see that when a donor is commended for following the 2001 Recommendation Against Aid-Tying, the previous positive relationship vanishes. Specifically, the coefficient for aid-tying becomes negative in Models 5–7. While this result is not statistically significant—likely due to the small number of "praising" mentions—it contrasts sharply with the strong positive relationship found in all other mentions. Neutral or non-praising mentions continue to reflect a donor's actual tying practices (Models 8–10). In substantive terms, a one-standard deviation increase in tied aid is related to an increase in the share of non-praising mentions in the summary section of the peer-review document of 4.4 percentage points—or 33.8 percent of its standard deviation ($p<0.001$). These findings are robust even when controlling for tied aid mentions in annual DAC reports.

Finally, we consider the possible effect of peer examiners. As discussed, DAC members who serve as examiners might have their own aid-giving practices in mind while assessing those of the donor under review. If an examiner country's own aid-tying track record is poor, it might be more lenient when assessing other donors' aid-tying practices. Conversely, when donors with excellent compliance with the 2001 Recommendation serve as examiners, they might be stricter enforcers of that standard. In short, this dynamic would generate the opposite relationship between

Table 2 Effect of reviewed-donor practice on tied aid evaluation: mention sentiment

	Praising mentions			Other mentions		
	Full document	Report	Summary	Full document	Report	Summary
	(5)	(6)	(7)	(8)	(9)	(10)
Tied Aid $_{t-1:t}$	−0.005	−0.003	−0.013	0.132***	0.121***	0.195***
	(0.014)	(0.012)	(0.039)	(0.023)	(0.022)	(0.053)
Mentions in DCR $_{t-1}$	3.742	1.794	10.600	9.287	7.514	12.167
	(4.484)	(3.897)	(12.722)	(7.656)	(7.278)	(17.525)
Constant	0.029***	0.029***	0.040*	0.026*	0.025*	0.045+
	(0.007)	(0.006)	(0.019)	(0.011)	(0.011)	(0.026)
Observations	116	116	116	116	116	116
R^2	0.007	0.002	0.007	0.232	0.217	0.111
Adjusted R^2	−0.010	−0.015	−0.011	0.219	0.203	0.095

† $p<0.1$; * $p<0.05$; ** $p<0.01$; *** $p<0.001$

examiner aid-tying practice and tied aid mentions than what we see with the reviewed donor's practice: as examiner aid-tying practice increases, the prevalence of tied aid discussion and the likelihood of receiving a recommendation against aid tying should decrease.

At the same time, the DAC peer review process is codified and organized driven by the Secretariat. Representatives of the examiner countries participate in review missions to the reviewed donor's capital and their main aid agencies. The reviewers then send their observations to a designated DAC team that drafts the Secretariat's Report. This team also draws on its own analysis of the reviewed donor's aid program. The two examiner countries participate in the peer review meeting, where the review documents are discussed and finalized. Given this process, we expect peer examiners to have a more limited ability to shape the content of the Secretariat's Report in a systematic way. Instead, it is more likely that they might have more influence in the front-end summary section of the review or in the inclusion of a specific recommendation against aid tying.

Table 3 presents a third set of linear regressions. This time, in addition to the aid-tying practice of the donor under review, we assess the effect of the aid-tying practice of the reviewee's examiners. In our main specification, examiner aid-tying practice is measured as the average of the two-year tied aid share across both reviewers, matching the corresponding measure for the reviewee. If data for either of the two examiners are missing, we use the practice of the other. Again, these models show a positive and significant effect of the reviewed donor's aid-tying track record on tied aid mentions in the review. However, the coefficients for the effect of examiner practice and its interaction with reviewed donor practice are not statistically significant, providing little support for the notion that examiner countries purposefully influence DAC peer reviews based on their own compliance with the organization's standards. These results are qualitatively similar when we consider alternative lag structures, including three-year averages (Table A6) and a one-year lag (Table A7). While acknowledging these null results, it is perhaps still interesting to note that the coefficients for examiner aid-tying practice are positive for the peer review document

Table 3 Effect of reviewed donor practice on tied aid evaluation: peer examiners

	Tied aid mentions			
	Full document	Report	Summary	Recommendation
	(11)	(12)	(13)	(14)
Tied Aid $_{t-1:t}$	0.142*	0.124*	0.269+	0.838*
	(0.056)	(0.051)	(0.143)	(0.382)
Avg. Examiner Practice	−0.008	0.000	−0.044	0.064
	(0.061)	(0.055)	(0.155)	(0.457)
Mentions in DCR $_{t-1}$	13.057	9.410	22.532	86.277
	(10.235)	(9.355)	(26.110)	(64.685)
Constant	0.055**	0.053**	0.087+	0.123
	(0.018)	(0.016)	(0.045)	(0.115)
Observations	115	115	115	104
R^2	0.150	0.146	0.066	0.099
Adjusted R^2	0.119	0.115	0.032	0.062

† $p<0.1$; * $p<0.05$; ** $p<0.01$; *** $p<0.001$

as a whole and its report section, while negative for its summary section and for the presence of a recommendation against aid tying. The latter two are indeed the places where, as discussed, it is more sensible to expect that examiners might be able to impact review documents.

In the appendix, we probe alternative ways of aggregating the peer reviewers' measures of aid-tying. Specifically, we generate dichotomous indicators for whether (1) at least one examiner is above the DAC average (Table A8), and (2) at least one examiner is above the 75th percentile of the DAC average—capturing the presence of the worst offender on tied aid (Table A9). The results warrant no change in our conclusions that peer reviewers' own aid tying practice has no consistent relationship with mentioning of aid tying in the peer review report.

Taken together, the results of these three sets of analyses point to a robust relationship between the reviewed donor's track record on tied aid and the content of its evaluation by the DAC: donors that go against the IO's standards by tying greater proportions of their aid receive reviews with a more prevalent discussion of tied aid and are more likely to get an explicit recommendation against aid tying. While the expectation behind this relationship is intuitive, evidence of a strong link between donor practice and its evaluation in peer review is a first sign that the organization's assessment process is working as intended, thus providing peer-to-peer accountability in foreign aid. The lack of evidence for a link between the practice of peer examiner states and review content further indicates that the DAC's peer review system assesses its members according to the same criteria, and does not leave room for its members' particularistic interests to influence its outcomes. This standardized assessment also extends to newer DAC members, such as South Korea and Central and Eastern European donors. Robustness checks confirm that these nations are not granted extra slack; rather, they exhibit the exact same robust relationship between tied aid practices and evaluative peer review mentions as established donors (Table A10).

5 Conclusion

We examine to what extent peer review among donor governments is able to fulfill its role in ensuring accountability against collective norms of good donorship. Using the policy norm of aid tying—a practice that can undermine aid effectiveness by forcing recipients to procure goods and services from the donor country—we study whether the aid-tying practice of a given donor affects the likelihood of a peer review to include a recommendation against such practice. We exploit a newly collected corpus of review documents covering all 143 peer reviews of the 32 OECD DAC donors from 1994 to 2020, combined with data on aid-tying practice from the Creditor Reporting System.

The findings suggest that DAC peer reviews interpret donor practices in ways that are broadly consistent with observable behavior and shared standards. We find a positive and statistically significant relationship between a donor's actual aid-tying practice and whether its review contains more mentions of tied aid or a recommendation against aid tying. Distinguishing between praising, neutral, and negative mentions of the tied aid topic in peer reviews, we further show that this positive relationship is indeed stronger when we remove praising mentions from our analyses. There is no evidence to support the notion that other aspects, like donor age or a peer examiner's track record, would translate into biased review content outcomes.

Before discussing the wider implications of our findings, we note the limitations. We only studied one peer review mechanism—the OECD DAC peer review—leaving the question of generalizability unaddressed. We chose the DAC peer review mechanism because of its prominence in shaping norms of good donorship, as well as the level of transparency that allowed us to encode data from the entire universe of peer reviews. Moreover, with the return of great power politics and national interest in development cooperation (Blair et al., 2022; Anyiam-Osigwe et al., 2026a and b; Bau & Dietrich, 2026; Bau et al., 2026a; Carcelli, 2023; Gulrajani, 2023), the issue of aid tying will gain further prominence in the near future, thus increasing the importance of studying aid tying. Another limitation is that our analyses are correlational. While the focus of our paper has been to examine whether aid practice affects discourse, the opposite mechanism whereby aid discourse affects aid practice is equally plausible and worthy of investigation. We encourage future research to examine this issue, as it holds promise to inform our understanding of whether, when, and how international norms shape state behavior. Our findings provide grounds for cautious optimism, suggesting that DAC peer reviews can fulfill their intended accountability function. Credible and unbiased evaluations of member performance are a key precondition for peer-to-peer accountability in international organizations. In the context of development cooperation, DAC peer reviews monitor donor compliance with agreed standards for effective aid giving, including the 2001 Recommendation on Untying Official Development Assistance. Our finding that the content of DAC peer reviews reflects the aid-tying practices of the reviewed donor provides new evidence that this mechanism produces credible assessments of donor performance and supports fair evaluation of compliance with shared aid commitments. We thereby complement previous work showing that crucial peer review decisions, including the choice of peer examiner countries and the aid recipients that are examined in-depth through field visits, largely follow DAC guidelines without much room for donor particularistic interest to influence the process (Iannantuoni et al., 2025).

Our results invite theorizing on the scope conditions for when and why peer reviews are effective in promoting accountability and enhancing compliance among member states. From a theoretical perspective, the DAC peer review mechanism—due to its active involvement of the DAC secretariat and its requirement of consensus on the final peer review report—is relatively more insulated against attempts by individual donors to advance their particularistic interests than other peer review mechanisms with more limited institutional safeguards (Jongen, 2018). While our empirical analysis focused on aid tying as a policy norm that touches on salient interests of member states, the technocratic nature of development policy and DAC decision-making may lead us to expect that member states might be more willing to interfere in the content of peer reviews in non-DAC venues where public scrutiny is higher and decisions tend to be more politicized given their heterogeneous membership, like the UN Human Rights Council (Steinert & Lang, 2025). Yet, we may expect our results for the DAC peer review to better generalize to other peer review mechanisms within the OECD universe.[17] The fundamental architecture of these peer-review mechanisms—a process led by reviewer states with strong support from the Secretariat as well as an emphasis on policy learning rather than legal sanctions—suggests that our findings would likely remain consistent. To complement the findings of our paper, we welcome further research into the varieties of peer review systems that work to reduce accountability deficits in global governance in different ways and across different issue areas.

From a policy perspective, our results suggest a need to uphold the collective commitments toward the DAC peer review mechanism and specifically its continued assessment of aid tying practices. With the renewed focus on national interests in aid allocation, there is a risk that tied aid once again becomes a salient issue capable of incentivizing donor interference in DAC peer evaluation. Aid tying remains a contentious topic, particularly among newer DAC members that have tied a disproportionate share of their assistance since joining. Moreover, a recent resurgence of aid tying among some established donors suggests that the longstanding consensus may be weakening. In this context, the DAC peer review serves as a vital institutional mechanism. By generating information on aid-tying practices, it ensures continued visibility and debate—either reinforcing the existing norm or creating opportunities to revise it (Hale, 2017). Additionally, revelations of malpractice can sound the alarm and mobilize domestic actors, such as development NGOs that oppose aid tying (Chaudoin, 2014).

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s11558-026-09639-1.

Acknowledgements We thank the three guest editors (Stephen Chaudoin, Christina Davis, and Christoph Mikulaschek) and the anonymous reviewers for excellent feedback. Thanks to Çağlayan Başer, Ekrem Başer, Sarah Bermeo, Shannon Carcelli, Stephen Chaudoin, Xinyuan Dai, Brian Gaines, Desh Girod, Chris Grady, Julia Gray, Magnus Lundgren, Helen Milner, Theresa Squatrito, Seiki Tanaka, Oliver Westerwinter, and Matt Winters for feedback and useful conversations. Thank you also to Kristina Enger and Tomas Spahni for invaluable research assistance.

Author contributions Author contributions to research design and conceptualization: S.D. (33%), A.I. (33%), B.R. (33%); data collection: A.I. (55%), B.R. (45%); statistical analysis: A.I. (30%), B.R. (70%); writing: S.D. (33%), A.I. (33%), B.R. (33%); project management: S.D. (100%). The order of authors does not reflect the significance of authors' contributions.

Funding Open access funding provided by University of Geneva. This research was supported by the Swiss National Science Foundation.

[17] The OECD peer review under the Economic and Development Review Committee is a highly structured, adversarial, yet collaborative peer review process, for which our findings will likely travel. Like DAC peer review, the process begins with the OECD Secretariat drafting a comprehensive report on a country's economy based on data and on-site missions. Then representatives from all member governments gather in Paris, and the reviewee country must formally defend its policies. Two peer countries are assigned as "lead examiners" to question the reviewed country's economic strategy and choices. After intense debate, the draft report is revised. Crucially, the final OECD Economic Survey must be approved by unanimous consensus—meaning the country being reviewed must ultimately agree to the wording of the critique and recommendations (Conzelmann, 2014).

Data availability Replication data and the supplemental appendix for this article are freely available on Harvard Dataverse (https://dataverse.harvard.edu/previewurl.xhtml?token=504f5fc5-07f4-4542-820b-390029469ae9).

Declarations

Conflict of interest The authors declare no conflict of interest.

References

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

Annen, K., & Kosempel, S. (2009). Foreign aid, donor fragmentation, and economic growth. *The B.E. Journal of Macroeconomics, 9*(1), 1–30.

Anyiam-Osigwe, T., & Vreeland, J. (2026). Japan and the African Development Bank. Forthcoming at *International Studies Quarterly*.

Anyiam-Osigwe, T., Dietrich, S., & Zeitz, A. (2026a). Citizen perceptions of the EU's Global Gateway and China's Belt and Roadway Initiative: Evidence from South Africa and Nigeria." Forthcoming at *Journal of European Integration*.

Anyiam-Osigwe, T., Dietrich, S., & Zeitz, A. (2026b). Preferred Partners?: Co-financing with the African Development Bank and the legitimation of foreign assistance. AidData Working Paper, 142.

Ashoff, G. (2013). 50 Years of peer reviews by the OECD's Development Assistance Committee: An instrument of quality assurance and mutual learning. *German Development Institute*.

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

Bau, N., & Dietrich, S. (2025). What geopolitical returns does ODA bring? *Kiel Institute Working Paper 2305*.

Bau, N., & Dietrich, S. (2026). Guaranteeing the Global Gateway: National development finance institutions and the transformation of EU development finance. *Forthcoming at Journal of European Integration*.

Bau, N., Dietrich, S., Fleiner, K., & Iannantuoni, A. (2025). Examining new donors in the OECD's Development Assistance Committee. *Journal of International Development*. https://doi.org/10.1002/jid.70000

Bau, N., Dietrich, S., Fleiner, K., & Iannantuoni, A. (2026b). Populism and the politics of aid classification: Evidence from the OECD DAC. Forthcoming at *International Studies Quarterly.*

Bau, N., Dietrich, S., Qian, J., & Trinh, D. (2026a). The institutional politics of development finance: ODA's role in shaping subnational DFI investments. Forthcoming at *Review of International Organizations*.

Biermann, F., & Siebenhüner, B. (Eds.). (2009). *Managers of global change: The role and relevance of international bureaucracies*. The MIT Press.

Blair, R., Marty, R., & Roessler, P. (2022). Foreign aid and soft power: Great power competition in Africa in the early twenty-first century. *British Journal of Political Science, 2*, 1355–1376.

Breen, M., & Doak, E. (2021). The IMF as a global monitor: Surveillance, information, and financial markets. *Review of International Political Economy, 30*(1), 307–331.

Broome, A., Homolar, A., & Kranke, M. (2018). Bad science: International organizations and the indirect power of global benchmarking. *European Journal of International Relations, 24*(3), 514–539.

Busch, P.-O., & Liese, A. (2016). The authority of international public administrations. *International bureaucracy: Challenges and lessons for public administration research* (pp. 97–122). Springer.

Carcelli, S. (2023). Special interests in foreign policy bureaucracies: Evidence from foreign aid. *Journal of Politics, 85*(3), 905–918.

Carcelli, S. P. (2024). Bureaucratic structure and compliance with international agreements. *American Journal of Political Science, 68*(1), 177–192.

Carnegie, A., Clark, R., & Zucker, N. (2024). Global governance under populism: The challenge of information suppression. *World Politics, 76*(4), 639–666.

Carraro, V., Conzelmann, T., & Jongen, H. (2019). Fears of peers? Explaining peer and public shaming in global governance. *Cooperation and Conflict, 54*(3), 335–355.

Chaudoin, S. (2014). Audience features and the strategic timing of trade disputes. *International Organization, 68*(4), 877–911.

Conzelmann, T. (2014). The politics of peer reviewing: Comparing the OECD and the EU. In T. Blom, & S. Vanhoonacker (Eds.), *The politics of information. The case of the European Union.* Chapter 4 (pp. 49–62).

Davis, C. L. (2023). *Discriminatory clubs: The geopolitics of international organizations*. Princeton University Press.

Dietrich, S. (2021). *States, markets, and foreign aid*. Cambridge University Press.

Dietrich, S., Iannantuoni, A., & Winters, M. (2025). Which countries does the OECD Development Assistance Committee (DAC) talk about when it talks about international development? *World Development*. https://doi.org/10.1016/j.worlddev.2025.107195

Dreher, A., Sturm, J.-E., & Vreeland, J. R. (2009). Development aid and international politics: Does membership on the UN Security Council influence World Bank decisions? *Journal of Development Economics, 88*(1), 1–18.

Eurodad. (2021). Strings still attached: Unmet commitments on tied aid. https://www.eurodad.org/unmet_commitments_tied_aid. Accessed: 2021–11–18.

Eurodad. (2024). Exposing tied aid: Preventing donor countries from getting rich on their own aid. https://www.eurodad.org/exposing_tied_aid_2024. Accessed: 2024–04–04

Farrell, H., & Newman, A. L. (2019). Weaponized interdependence: How global economic networks shape state coercion. *International Security, 44*(1), 42–79.

Ganga, P. (2025). Ties that bind: The impact of tied aid on development. *Studies in Comparative and International Development, 60*(3), 613–631.

Groenendijk, N. (2009). EU and OECD benchmarking and peer review compared. *Third EUCE Annual Conference: The EU in a Comparative Perspective*, 1–18.

Gui, T. (2024). Behind closed doors: Informal influence on United Nations staffing and pathologies of international bureaucracies. *Global Policy, 15*(4), 615–630.

Gulrajani, N. (2023). Bilateral donors and the age of the national interest: What prospects for challenge by development agencies? *World Development, 96*, 375–389.

Hai, Z. (2025). The global politics of scientific consensus: Evidence from the Intergovernmental Panel on Climate Change. *International Organization, 79*(2), 233–256.

Hale, T. N. (2008). Transparency, accountability, and global governance. URL: http://www.jstor.org/stable/27800692

Hale, T. N. (2017). Under what conditions does international review alter national policy? Refining concepts and building theory. *Political Economy of International Organization*.

Hollyer, J. R., Rosendorff, B. P., & Vreeland, J. R. (2014). Measuring transparency. *Political Analysis, 22*(4), 413–434.

Honig, D., & Weaver, C. (2019). A race to the top? The aid transparency index and the social power of global performance indicators. *International Organization, 73*(3), 579–610.

Iannantuoni, A., Dietrich, S., & Reinsberg, B. (2025). Who reviews whom, where and why? Evidence from the peer review process in the OECD Development Assistance Committee. *International Studies Quarterly, 69*(1), sqae138.

Jongen, H. (2018). The authority of peer reviews among states in the global governance of corruption. *Review of International Political Economy, 25*(6), 909–935.

Jongen, H. (2021). Peer review and compliance with international anti-corruption norms: Insights from the OECD working group on bribery. *Review of International Studies, 47*(3), 331–352.

Jorgens, H., Kolleck, N., Saerbeck, B., & Well, M. (2016). Orchestrating (bio-) diversity: The secretariat of the Convention of Biological Diversity as an attention-seeking bureaucracy. *International bureaucracy: Challenges and lessons for public administration research* (pp. 73–95). Springer.

Keesing, D. B. (1998). *Improving trade policy reviews in the world trade organization* (Vol. 52). Peterson Institute.

Kelley, J. G., & Simmons, B. A. (2021). Governance by other means: Rankings as regulatory systems. *International Theory, 13*(1), 169–178.

Keohane, R. O. (1984). *After hegemony: Cooperation and discord in the world political economy*. Princeton University Press.

Kilby, C. (2011). Informal influence in the Asian Development Bank. *Review of International Organizations, 6*(3), 223–257.

Koremenos, B., Lipson, C., & Snidal, D. (2001). The rational design of international institutions. *International Organization, 55*(4), 761–799.

Lim, D. Y. M., & Vreeland, J. R. (2013). Regional organizations and international politics: Japanese influence over the Asian Development Bank and the UN Security Council. *World Politics, 65*(1), 34–72.

McLean, E., & Park, J. (2025). Containing China's rising power in international organizations: Earmarked funding and influence in multilateral development banks. *Review of International Organizations.*

Mikulaschek, C. (2023). The responsive public: How European Union decisions shape public opinion on salient policies. *European Union Politics, 24*(4), 645–665.

Mitchell, S. M., & Hensel, P. R. (2007). International institutions, cooperation, and compliance with agreements. *American Journal of Political Science, 51*(4), 721–737.

Morse, J. C. (2019). Blacklists, market enforcement, and the global regime to combat terrorist financing. *International Organization, 73*(2), 511–545.

Novosad, P., & Werker, E. (2019). Who runs the international system? Nationality and leadership in the United Nations secretariat. *Review of International Organizations, 14*(1), 1–33.

OECD. (1965). DAC recommendation on assistance and financing targets. Technical report. Organization for Economic Cooperation and Development – Donor Assistance Committee.

Oksamytna, K. (2018). Policy entrepreneurship by international bureaucracies: The evolution of public information in UN peacekeeping. *International Peacekeeping, 25*(1), 79–104.

Rabinowitz, G. (2024). The UK gets an 'easy ride' from the OECD-DAC. *Bond Report*.

Reinsberg, B., & Heinzel, M. (2026). Siloes of influence: The political economy of resourcing international organizations. *Review of International Organizations*. https://doi.org/10.1007/s11558-026-09615-9)

Sadeh, T., & Porath, Y. (2020). Autonomous agencies and relational contracts in government bond issues. *Regulation & Governance, 14*(4), 741–763.

Simmons, B. A. (2000). International law and state behavior: Commitment and compliance in international monetary affairs. *American Political Science Review, 94*(4), 819–835.

Steinert, C., & Lang, V. (2025). The illiberal challenge to the liberal international order: Evidence from the UN Human Rights Council. *Working Paper*.

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

Terman, R., & Voeten, E. (2018). The relational politics of shame: Evidence from the Universal Periodic Review. *Review of International Organizations, 13*(1), 1–23.

Tingley, D., & Tomz, M. (2022). The effects of naming and shaming on public support for compliance with international agreements: An experimental analysis of the Paris Agreement. *International Organization, 76*(2), 445–468.

Voeten, E. (2001). Outside options and the logic of Security Council action. *American Political Science Review, 95*(4), 845–858.

Von Bogdandy, A., & Goldmann, M. (2008). The exercise of international public authority through national policy assessment. *International Organization Law Review, 5*, 241–298.

Zaccaria, G., & Reinsberg, B. (2025). Informal state influence in international organizations: Examining the link between executive head nationality and earmarked funding. *European Journal of International Relations*. https://doi.org/10.1177/13540661251355045

Zhang, X., & Chen, Y. (2025). Emerging power, tight wallet: How China chooses where to earmark funds in the UN. *Review of International Organizations.*