Making Global Governance Accountable: Civil Society, States, and the Politics of Reform

Ranjit Lall   University of Oxford

Abstract: *Recent decades have witnessed the adoption of unprecedentedly broad and inclusive accountability mechanisms by many major international institutions, from grievance redress systems to transparency policies. What explains the establishment of these mechanisms—and why have only some institutions embraced them? I argue that adoption is more likely when member states, in particular the most powerful, face "bottom-up" pressures for accountability from dense transnational civil society networks—networks with the capacity to build leverage through agenda setting, coalition building, and advocacy strategies—and when institutions perform governance tasks that are costly to monitor. Analysis of a rich new dataset shows that adoption is positively related to the density of international nongovernmental organizations in an institution's issue area—including only those based in powerful member countries—and that this relationship is stronger when governance tasks entail high monitoring costs. Statistical tests are complemented by qualitative evidence from interviews and other primary sources.*

Verification Materials: The data and materials required to verify the computational reproducibility of the results, procedures, and analyses in this article are available on the American Journal of Political Science Dataverse within the Harvard Dataverse Network, at: https://doi.org/10.7910/DVN/R4USRU.

As international institutions have grown in number, scope, and authority in recent decades, they have faced growing calls to become more accountable, especially to those beyond their membership—that is, to provide broader opportunities to understand, evaluate, and influence their actions and to discipline them if they fall short of expectations (Grant and Keohane 2005). In response, many institutions have established what might be termed *multistakeholder* accountability (MSA) mechanisms: formalized routines and procedures, based broadly on principles of democratic governance, that seek to enhance the capacity of diverse public and private stakeholders to monitor, assess, and shape institutional activities. Ranging from evaluation offices and grievance redress systems to access-to-information policies and participatory governance arrangements, MSA reforms have ushered in a new era of transparency, inclusiveness, and responsive-

ness in global governance—an era where, in the eyes of some observers, the separation between domestic and international modes of political organization is rapidly eroding (Kingsbury, Krisch, and Stewart 2005).

The adoption of MSA mechanisms, however, has been far from universal. Small and low-profile institutions have largely maintained the traditional state-centric model of global public accountability premised on participation in institutional governing bodies. Even among large and high-profile institutions whose activities are subject to intense scrutiny by states, the range and depth of MSA reforms has varied substantially. The divergent trajectories of the World Bank, the International Monetary Fund (IMF), and the World Trade Organization (WTO)—often regarded as the three central pillars of global economic governance—are a case in point. Over the past three decades, the World Bank has established stringent versions of almost every type of MSA

Ranjit Lall, Associate Professor, Department of Politics and International Relations, University of Oxford, Manor Road Building, Manor Road, Oxford OX1 3UQ, United Kingdom (ranjit.lall@politics.ox.ac.uk).

mechanism and was the first intergovernmental organization (IGO) to establish an access-to-information policy, a grievance redress system (the famous Inspection Panel), and an independent evaluation office. The IMF, in contrast, possesses a strong evaluation mechanism but a weak access-to-information policy and no grievance redress procedure. The WTO has been even more resistant to reform, adopting a limited access-to-information policy shortly after its creation and failing to establish any kind of evaluation or grievance redress function since then.

What explains the creation of MSA mechanisms—and why have only some of the most prominent and scrutinized institutions of global governance embraced them? The answer matters from both a normative and a practical standpoint. Accountability is widely regarded as a desirable characteristic for public institutions, both in its own right—those who wield political authority, many believe, should be subject to the scrutiny, judgment, and sanction of those who delegated and are affected by such power—and for its close connection to legitimacy, a key determinant of stakeholder compliance with international rules (Grant and Keohane 2005). Furthermore, a growing body of empirical evidence indicates that MSA mechanisms can improve governance outcomes by incentivizing closer bureaucratic adherence to institutional policies and protocols (Buntaine 2015; Honig, Lall, and Parks 2022; Lall 2023; Nielson and Tierney 2003).

Previous scholarship on the origins of MSA mechanisms mostly falls into one of two camps. The first camp undertakes case studies of specific reform episodes, in some instances drawing on personal accounts of those involved in campaigns for accountability (Fox and Brown 1998; Keck and Sikkink 1998; Scholte 2011). A recurring theme is the critical role of "bottom-up" pressures from transnational civil society (TCS) in instigating and driving reform, often in the face of entrenched institutional resistance. While offering valuable insights into reform dynamics, these studies do not seek to identify common traits of successful TCS campaigns or to explain why the strength of MSA mechanisms differs across institutions. The second camp investigates the determinants of reform with systematic empirical data. This set of analyses employs cross-sectional or random-effects panel designs to examine the adoption of varying subsets of MSA mechanisms, including transparency policies (Ecker-Ehrhardt 2018; Grigorescu 2007), internal oversight mechanisms (Grigorescu 2007; 2010), and access to institutional bodies (Sommerer and Tallberg 2019; Tallberg et al. 2013). On the explanatory front, they focus primarily on "top-down" political and strategic factors affecting member states' demand for reform, such as their regime type, foreign policy preferences, and group size.

This study proposes a general explanation for differences in the adoption of MSA mechanisms that contributes to a third, smaller strand of literature emphasizing the interplay between bottom-up and top-down forces (Pallas and Uhlin 2014; Pallas and Urpelainen 2012; Raustiala 1997). Building on structural theories of institutional change, I argue that the strength and effectiveness of bottom-up pressures for MSA reforms are a function of the density of TCS organizations in an institution's issue area, in particular those based in powerful member countries. High levels of TCS density facilitate agenda setting, coalition building, and advocacy strategies by activists, enabling them to build and sustain political leverage over states of varying sizes and capabilities. The impact of bottom-up pressures is not uniform across institutions, however. Heeding principal-agent analyses of delegation, I posit that states' incentives to support TCS reform demands depend on the ease with which they can monitor the exercise of delegated governance tasks, which determines the benefits they, as "principals," derive from new opportunities to acquire information on, evaluate, and control bureaucratic "agents." In short, my approach calls attention to a key structural determinant of TCS influence and a moderating strategic feature of institutional context to which previous approaches—whether bottom-up, top-down, or hybrid—have paid little systematic attention.

I subject the argument to statistical tests based on a rich original dataset covering five types of MSA mechanisms—*transparency*, *evaluation*, *redress*, *investigation*, and *participation* mechanisms—across 52 international institutions between 1960 and 2018. Unlike previous analyses of MSA reforms, I employ a fixed-effects design that controls for sources of unobserved institution- and time-specific heterogeneity. I find that the adoption of MSA mechanisms is positively related to the density of international nongovernmental organizations (INGOs) in an institution's issue area, and that this association is robust to a variety of controls, samples, and estimation strategies, including the use of instrumental variables (IV) to deal with possible reverse causation. Approximately two-thirds of the relationship is explained by INGOs headquartered in member states and half by INGOs headquartered solely in the five most powerful of these countries. In addition, I show that issue-specific INGO density is more strongly associated with adoption when institutions perform governance tasks with high monitoring costs, such as implementing operational activities in the field, than when they exercise tasks that are easier to oversee, such as facilitating

interstate agreements. Finally, I furnish evidence for a variety of more subtle implications of my argument, including that higher TCS density is attended by increased publicity for MSA campaigns and that pressures on powerful member states are more likely to bring about reform when such countries enjoy greater influence in institutional decisionmaking.

The statistical analysis is complemented by qualitative evidence on a range of notable cases and patterns in the data, which draws on key informant interviews and other primary and secondary sources. The goal here is not to enhance the robustness of the statistical results but to illustrate the plausibility of posited causal processes that are difficult to probe quantitatively, in particular those concerning TCS's agenda-setting and lobbying activities. In doing so, the qualitative examination sheds light on the puzzling discrepancy in the strength of MSA mechanisms adopted by the World Bank, the IMF, and the WTO, tracing how differences both in the density of INGOs in their respective issue areas—high in economic development, moderate in finance, and low in trade— and in the costs of monitoring their respective tasks— high for development assistance, high for financial stabilization, and low for facilitating trade agreements—have resulted in widely varying propensities for reform.

The Rise of MSA Mechanisms: An Overview

MSA mechanisms can be traced back to experiments with self-evaluation and the inclusion of nonstate actors in governance processes around five decades ago, which laid the groundwork for more routinized systems for promoting transparency, widening participation, and addressing stakeholder grievances in the 1990s and 2000s. To systematically map the spread of these innovations, I collected time-series data on MSA structures in the 52 institutions included in official multilateral performance assessments conducted by donor governments over the past two decades (Lall 2017), a list of which is provided in the appendix. I selected this sample for three reasons. First, since the main criteria for inclusion in the assessments are funding levels and relevance to the donor's foreign policy priorities, the sample encompasses a substantial slice of the population of interest: large and politically significant institutions subject to close state scrutiny. Second, as documented in Appendix B in the online supporting information, the distribution of issue areas in the sample is similar to that in the wider population of IGOs. Third, many of the performance assess-

ments rate institutions on dimensions of accountability as well as performance. As these ratings are informed by the views of citizens, civil society groups, international bureaucrats, and academics in both developing and developed countries, they offer an opportunity to crosscheck my data against the perceptions of a variety of key stakeholders.

The dataset encompasses the five principal types of MSA mechanisms identified in previous studies and evaluations of accountability in global governance (including the One World Trust's Global Accountability Report): transparency, evaluation, redress, investigation, and participation mechanisms. The strength of each type of mechanism is captured by a 5-point index for every year from 1960 (or an institution's founding date) to 2018. The main data sources, enumerated for each institution in Appendix A in the online supporting information, are policy and governance documents (obtained in some cases from institutional archives), online reporting, and personal communications with institutions. The five indices can be summarized as follows (see Table 1 for coding rules):

1. *Transparency* is an additive index measuring whether institutions possess an access-to-information policy—a policy that enshrines the public's right to request (nonsensitive) information from them—and whether this policy guarantees automatic and timely disclosure and includes an independently managed appeals process for rejected disclosure requests (i.e., a process not managed by the secretariat itself).

2. *Evaluation* is an additive index measuring whether institutions possess an internal unit (e.g., office, department, division) responsible for monitoring and assessing their activities and whether this unit is organizationally independent, publicly discloses evaluation findings, and works with the secretariat to implement lessons for improving performance.

3. *Redress* is an additive index measuring whether institutions possess a mechanism for receiving, assessing, and addressing complaints from adversely affected external stakeholders and whether this mechanism is independently managed, guarantees confidentiality and nonretaliation for complainants, and includes systems for monitoring the implementation of remedial measures.

4. *Investigation* is an additive index measuring whether institutions possess a mechanism for investigating and sanctioning professional,

TABLE 1 Indices of Multistakeholder Accountability

Mechanism	Indicator	Description	Score
Transparency	Policy	Official access-to-information policy	1
	Presumption of disclosure	All information disclosed in absence of compelling reason to conceal	1
	Confidentiality	Sensitive types of information clearly defined	1
	Timeframe	Time frame for responding to information requests	1
	Appeals	Appeals process for rejected information requests	1
Evaluation	Unit	Unit responsible for evaluating performance	1
	Independence	Unit independent from secretariat	1
	Disclosure	Evaluation products publicly disclosed	1
	Response	Evaluations require response from management	1
	Follow-up	Implementation of evaluation recommendations monitored	1
Redress	Function	Mechanism for addressing stakeholder complaints	1
	Independence	Mechanism independent from secretariat	1
	Confidentiality	Confidentiality for complainants guaranteed	1
	Nonretaliation	Nonretaliation against complainants guaranteed	1
	Follow-up	Implementation of remedial measures monitored	1
Investigation	Function	Mechanism for investigating staff misconduct	1
	Independence	Mechanism independent from secretariat	1
	Confidentiality	Confidentiality for complainants guaranteed	1
	Nonretaliation	Nonretaliation against complainants guaranteed	1
	Follow-up	Implementation of remedial measures monitored	1
Participation	Access to governing body	External stakeholders are members of governing body	3
		External stakeholders participate in governing body	2
		External stakeholders observe/attend governing body	1
	Advisory body	External stakeholders represented in advisory body	1
	Consultation body	External stakeholders participate in consultation forum	1
Participation indicators multiplied by:	Unrestricted	Access granted to all external stakeholders	× 1
	Issue restrictions	Access restricted on basis of issue area	× 0.75
	Nonissue restrictions	Access restricted on basis of criteria other than issue area (e.g., expertise, financial contributions, location)	× 0.50
	Name restrictions	Access restricted to named or elected stakeholders	× 0.25

Notes: This table details the coding of five indices of multistakeholder accountability (MSA).

financial, or other misconduct by officials and whether this mechanism has the same four additional characteristics as *Redress* (i.e., independence, confidentiality, nonretaliation, and follow-up).

5. *Participation* is a multiplicative index measuring (1) the depth of access to policy processes granted to external stakeholders (such as civil society actors and corporations) and (2) the range of such stakeholders permitted access. Scores are averaged across three types of policy

organ: governing bodies, advisory councils, and consultation forums.[1]

[1]The index is defined by:

$$\text{Participation}_{a,b,c} = \sum_{a=1}^{A} \frac{\text{governing}_a \times \text{range}_a}{A}$$

$$+ \sum_{b=1}^{B} \frac{\text{advisory}_b \times \text{range}_b}{B}$$

$$+ \sum_{c=1}^{C} \frac{\text{consultation}_c \times \text{range}_c}{C}, \quad (1)$$

FIGURE 1 Multistakeholder Accountability Indices over Time
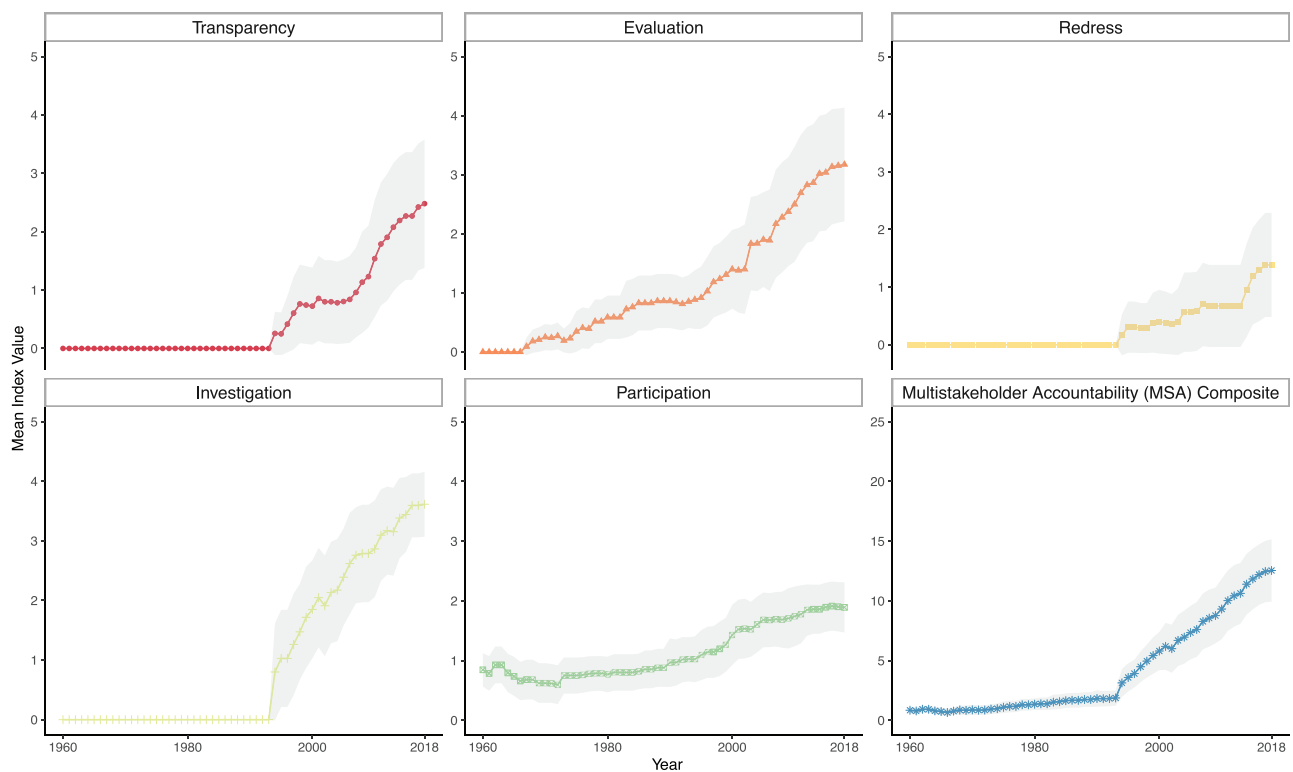
Notes: Mean values of the five multistakeholder accountability (MSA) indices defined in Table 1, plus their summation, between 1960 and 2018. Shaded regions represent standard deviation intervals.

The five indices are strongly associated: all 10 bivariate correlations are positive and statistically significant at the 1% level, with a mean *r* of 0.51. They are also highly correlated with almost 20 indicators of comparable dimensions of accountability in the donor-performance assessments mentioned earlier. While we should not expect these associations to be perfect—the assessments do not focus specifically on MSA mechanisms—they provide evidence that the indices are broadly consistent with how contemporary accountability structures are viewed by diverse stakeholders. All of the above correlations are reported in Appendix B in the online supporting information.

Figure 1 plots the mean value of each MSA index as well as a summative combination of all five (*MSA Composite*) over the full sample period. All indices exhibit a clear upward trend, albeit with some variation in timing. *Evaluation* begins its rise in the early 1970s,

where *governing$_a$*, *advisory$_b$*, and *consultation$_c$* are the depth of access to governing body *a*, advisory council *b*, and consultation forum *c*, respectively; *range$_a$*, *range$_b$*, and *range$_c$* are the breadth of access to such organs; and *A*, *B*, and *C* are the total number of such organs. This formula builds on the measurement of transnational access by Tallberg et al. (2013).

when several multilateral development banks (MDBs) and United Nations (UN) agencies—led by the UN Development Programme (UNDP) and the UN Environmental Programme (UNEP)—established internal assessment offices, and accelerated in the late 1990s as the paradigm of independent evaluation gained currency in foreign aid circles. *Participation* follows a similar but less pronounced trajectory, growing steadily during the 1970s as UNDP, UNEP, and other developmental and environmental institutions sought greater policy input from civil society (including in the governance of hosted treaties, beginning with the 1974 Convention on International Trade in Endangered Species of Wild Flora and Fauna), before receiving fresh impetus in the new millennium with the broader "opening up" of IGO bodies to nonstate actors (Tallberg et al. 2013). *Transparency*, *Redress*, and *Investigation* start their ascent in the mid-1990s, after a high-profile scandal involving a socially and environmentally destructive World Bank dam project on India's Narmada River culminated in the widespread adoption of access-to-information policies and stakeholder-complaint mechanisms—most prominently the Inspection Panel—by MDBs.

Importantly, as indicated by the widening standard-deviation intervals around most of the trend lines in Figure 1, the spread of MSA mechanisms has generally been accompanied by increasing variation in their strength across institutions. Figure 2 provides a disaggregated view of such variation—and the historical patterns described above—by plotting the six indices over time for all 52 institutions. Together, the figures indicate that while many major international institutions have introduced broader and more inclusive accountability mechanisms in recent decades, many others continue to rely heavily on the limited state-centric structures with which they were founded.

Bottom-Up Pressures and Top-Down Monitoring

MSA mechanisms entail distributional consequences for actors below, within, and above the state. New channels for monitoring, evaluating, and influencing international institutions facilitate oversight and control by states and external stakeholders, potentially increasing the gains of multilateral cooperation. At the same time, they impose costs on international bureaucrats, whose actions become easier to scrutinize, as well as on states that benefit from weak accountability structures. In light of these mixed implications, the impetus for accountability has typically come from outside rather than within institutions—specifically, from civil society associations with a comparably transnational scope. In this section, I identify a central driver of the strength and effectiveness of TCS pressures for MSA reforms—the density of INGOs operating in an institution's issue area, in particular, those located in powerful member states—and highlight how the impact of such pressures is moderated by the costs of monitoring governance tasks, a salient feature of institutional context that determines the gains states reap from new channels of oversight.

## The (Transnational Civil) Societal Roots of Reform

In close tandem with the emergence of MSA mechanisms has been another significant trend in world politics that has received much more attention: increasing engagement between international institutions and TCS, broadly defined as the space outside the formal political system where citizens undertake goal-oriented collective action across borders in varied forms including advocacy organizations, charitable foundations, social movements, faith groups, and labor unions. As the costs of international communications and travel have declined and states have become more interconnected in recent decades, TCS—particularly its most organized, resourced, and connected elements, which often take the form of INGOs—has emerged as a powerful political force at both the domestic and the international level. Leveraging a distinctive set of advocacy tactics, including lobbying campaigns targeting powerful actors, the dissemination of charged information, and the use of symbols, stories, and frames to widen issue resonance, TCS activists have shaped international policy outcomes in a range of areas (Keck and Sikkink 1998). Accountability practices are an important yet often overlooked example.

TCS influence over global accountability structures derives from a combination of agenda setting, coalition building, and advocacy capacities. Calls for MSA mechanisms tend to be triggered by scandals, crises, disasters, and other problems that expose poor institutional performance. Civic activists are often among the first to draw attention to these problems—they are sometimes directly involved in uncovering them—and in linking them to accountability deficits in the ensuing public narrative. Stronger accountability mechanisms are proposed as a practically and normatively desirable solution, providing a means of remedying performance failings while allowing stakeholders to more easily hold institutions responsible for their actions.

To build support for reform, TCS groups have leveraged their geographical scope and organizational skills to assemble broad-based coalitions of institutional stakeholders. Since reform typically requires the (implicit or explicit) approval of institutional governing bodies, member countries are particularly important allies (Pallas and Uhlin 2014). As noted earlier, however, concerns about the distributional effects of MSA mechanisms may deter states from endorsing TCS demands. Consequently, the advocacy tactics highlighted by scholarship on transnational activism—in particular lobbying, information dissemination, and publicity generation—are often critical for ensuring the successful passage of MSA reforms through the policy process.

Although almost all international institutions experience episodes of poor performance and have at some point faced calls for greater accountability, the intensity of bottom-up pressures for MSA reforms varies widely across policy domains. Some issues, such as economic development, public health, and human rights, are the focus of substantial and widespread civic activity, a pattern generally attributed to their perceived salience

FIGURE 2 Multistakeholder Accountability Indices for Individual Institutions, 1960–2018
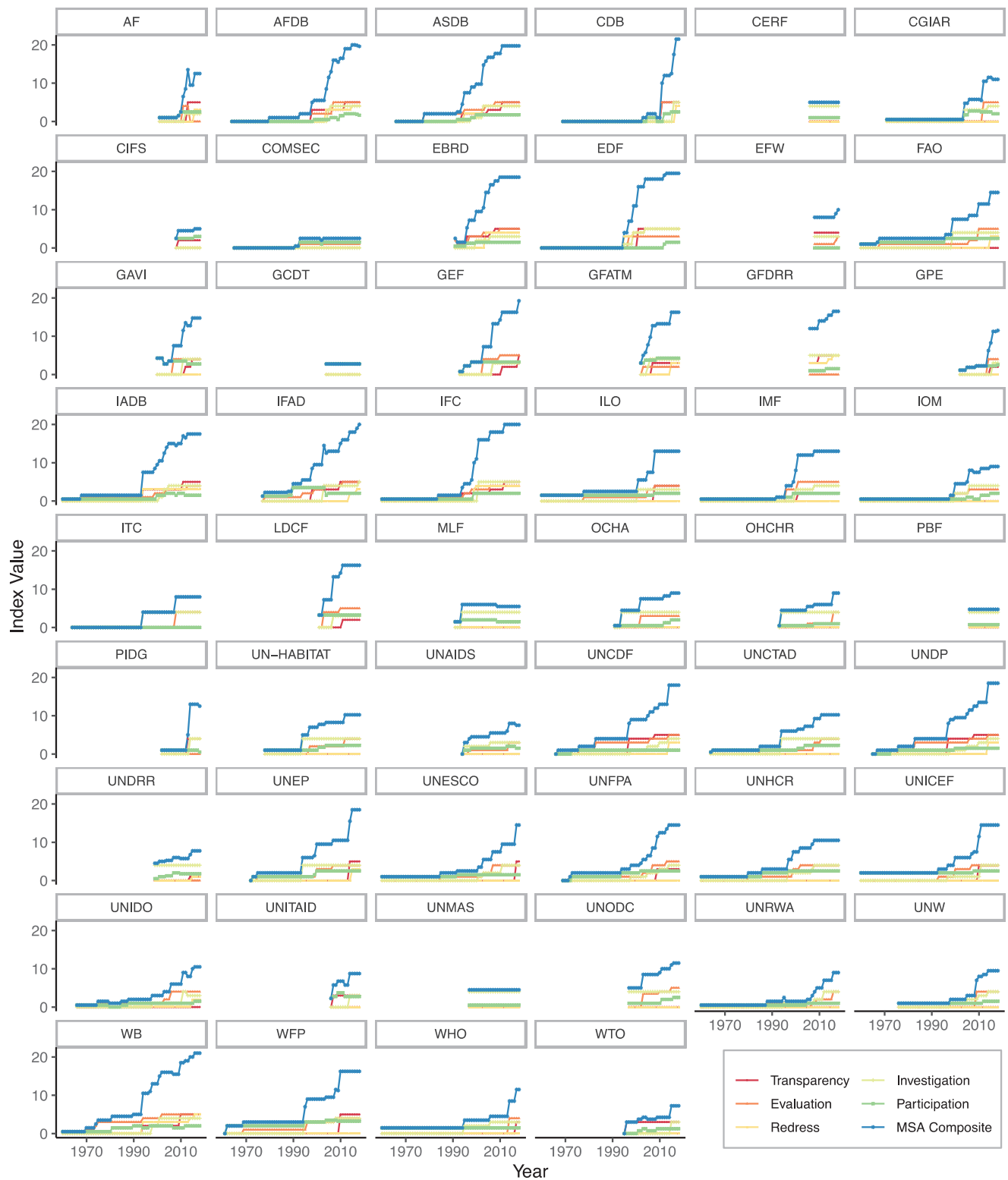
and urgency, centrality to human welfare, and resonance with diverse cultures and societies. In the words of Keck and Sikkink, they "speak to aspects of belief systems or life experiences that transcend a specific cultural or political context" (1998, 204). Other issues, such as competition policy, diplomacy, and transportation, tend to be perceived as less salient and pressing and to resonate less broadly across cultural and societal contexts, limiting the extent to which they inspire transnational collective action.

As highlighted by structural perspectives on institutional change, such as theories of organizational ecology (Abbott, Green, and Keohane 2016) and strategic choice (Jupille, Mattli, and Snidal 2013), population-level characteristics—most notably the density of actors addressing a particular issue—can profoundly affect patterns of behavior and influence within an institution's environment. In issue areas characterized by high levels of TCS density, case studies of successful transnational advocacy campaigns suggest, activists can more effectively deploy strategies for promoting policy reform. A denser constellation of TCS groups, in particular INGOs, should be more capable of detecting performance problems, bringing them to the attention of the wider stakeholder community, and fashioning a public discourse around the need for stronger accountability. Furthermore, it should lead to the formation of larger, more diverse, and better resourced coalitions for reform, which should be capable of mobilizing more intense and sustained pressure on member states via lobbying and information-dissemination activities (especially when comprising a high proportion of INGOs). The upshot should be broader and deeper support for MSA mechanisms both within and outside institutions.

The discussion thus far leads to the hypothesis:

> *H1*: The density of TCS organizations in an institution's issue area is positively related to its propensity to adopt MSA mechanisms.

When it comes to shaping state preferences regarding MSA mechanisms, of course, not all TCS groups are equally willing or capable. Case studies indicate that, even when they have a wide geographical presence, civil society associations tend to focus their efforts on the government of their home country, where their resources, members, and political connections are usually concentrated (Fox and Brown 1998; Pallas and Uhlin 2014; Scholte 2011). Nor, however, do all *states* enjoy the same ability to influence policy outcomes in international institutions. In some institutions, such as the World Bank and the IMF, larger economies enjoy higher voting shares in governing bodies, rendering

their consent essential for the passage of MSA reforms (Grigorescu 2010). Even when voting shares are equally distributed, powerful countries are known to routinely steer outcomes toward their preferences through the exercise of informal influence (for instance, lobbying secretariats or issuing threats to weaker states).

The precise number of powerful member states whose support is needed to secure the adoption of MSA mechanisms is difficult to specify ex ante, depending on how an institution's particular constellation of formal and informal governance arrangements disperse power among the membership. Although the backing of a single dominant state may sometimes be sufficient for reform—particularly when governance structures concentrate policy influence—case evidence suggests that INGOs frequently need to assemble a wider "coalition of the powerful" (Fox and Brown 1998; Scholte 2011). These considerations suggest an additional two-part hypothesis:

> *H2*: The density of TCS organizations that operate in an institution's issue area and are headquartered in its (1) member states and (2) most powerful member states only is positively related to its propensity to adopt MSA mechanisms.

The Moderating Role of Governance Tasks

While structural perspectives highlight how population-level factors can spur institutional transformation, principal-agent theories of delegation remind us that states' desire to prevent opportunistic behavior by international bureaucrats can place hard parameters on the scope of policy change (Hawkins et al. 2006; Nielson and Tierney 2003). When granting authority to international institutions, the principal-agent approach emphasizes, states aim to strike a balance between giving bureaucrats enough discretion to successfully fulfil governance tasks and maintaining sufficient control to prevent undesired outcomes. The information generated by MSA mechanisms facilitates the monitoring not only of state compliance with international commitments, as emphasized by previous studies (Pallas and Urpelainen 2012; Raustiala 1997), but also of *bureaucratic* performance in executing governance tasks. Accordingly, I argue that when bureaucratic monitoring costs are high, whether because tasks are difficult to scrutinize or require the application of

TABLE 2 Typology of Governance Tasks

Governance Task	Example	Monitoring Costs
Facilitating agreements	WTO provides a forum for states to negotiate international trade agreements	Low
Monitoring compliance	ILO requires states to regularly report on their implementation of ratified conventions	Low
Capacity building	UNDP transfers information, skills, technology, and equipment to developing countries	Low
Designing interventions	IMF attaches binding economic policy conditions to its financial assistance programs	High
Allocating resources	World Bank offers loans to low- and middle-income countries for development projects	High
Implementing operations	UNICEF delivers humanitarian assistance to children around the world	High

Notes: This table summarizes six common governance tasks delegated to international institutions.

technical expertise, states will possess stronger incentives to support TCS demands for MSA reforms.

Depending on the specific problem they were created to tackle, international institutions may be delegated a variety of governance tasks, whose monitoring costs vary widely for states. The following six tasks, summarized in Table 2, are particularly common:

1. *Facilitating agreements.* The task of providing a forum for states to interact, build consensus, and develop rules, standards, and norms tends to be straightforward for them to monitor. International bureaucrats are mainly required to perform hosting and convening functions, such as providing physical space, administrative assistance, and background information for meetings between delegates, who can directly observe whether and how well they exercise these functions. Institutions whose primary task is facilitating agreements, such as the WTO and the UN Conference on Trade and Development (UNCTAD), thus tend to have relatively small secretariats (based predominantly at their headquarters) and to be characterized as "member driven."

2. *Monitoring compliance.* The process of acquiring information on compliance with international agreements—traditionally regarded as one of the central functions of international institutions—is also, by its nature, readily ob-

servable by states. While some forms of compliance are harder for bureaucrats to observe and verify than others, "monitoring the monitors" rarely entails significant costs for states. Indeed, bureaucrats often lack the capacity or authority to monitor compliance themselves and thus rely heavily on self-reporting by states, which are directly privy to the information-gathering process. The International Labour Organization (ILO), for instance, monitors compliance with its conventions by asking states to submit regular updates on national implementation measures.

3. *Capacity building.* A common task among UN institutions is the transfer of information, skills, technology, equipment, and other kinds of human and physical capital to states for capacity-building purposes. This form of support, widely known as "technical assistance," is requested by and implemented jointly with recipient governments, allowing for comparatively easy oversight.

4. *Designing interventions.* Some institutions go beyond capacity building by directly shaping the content of domestic policies. This task tends to require more specialized knowledge and broader bureaucratic discretion, rendering it more difficult for states to monitor. A prominent example are the conditions attached to IMF loan programs, which are crafted on the basis of complex information about local economic

and political circumstances and sectoral policy expertise that recipient governments frequently lack (Martin 2006).

5. *Allocating resources.* The primary task of international financial institutions (including the IMF) is to pool and allocate material resources to states for economic purposes, usually via loans, credit lines, grants, or investments. This task is also costly for states to monitor: identifying and designing viable projects and programs to finance requires issue-specific expertise as well as detailed knowledge of the recipient country or entity (Martin 2006), restricting delegates to an often marginal role in the process. The Executive Boards of the World Bank and the IMF, for example, are regularly criticized for "rubber-stamping" proposals put forward by staff (Woods 2001, 87).

6. *Implementing operations.* The implementation of substantive operations in the field, such as peacekeeping and emergency relief missions, is inherently difficult to observe: key bureaucratic activities are dispersed across distant and frequently remote country offices and tend to require local information and organizational competences lacked by governments, creating a classic agency problem (Hawkins et al. 2006).

Two caveats about this list merit mention. First, it is not intended to be exhaustive. It excludes tasks exercised by few institutions, such as resolving disputes and authorizing sanctions, as well as essentially passive functions that institutions perform simply by virtue of existing, such as representing the international community and embodying norms. Second, even for a single task, these costs can vary across issues and over time. Some international trade and security agreements, for instance, have complex monitoring arrangements that cannot easily be overseen by states. On average, however, I expect the first three tasks to be less costly to monitor than the last three.

High monitoring costs deepen information asymmetries between states and international bureaucrats, expanding the scope for "agency slack." MSA mechanisms help to curtail such discretion by facilitating both the acquisition of information about bureaucratic performance and the sanctioning of undesired behavior and outcomes. In doing so, they should incentivize bureaucrats to enhance their effort and productivity both in anticipation and as a consequence of state sanctions (Buntaine 2015). As monitoring costs rise, therefore, MSA mechanisms should yield greater strategic benefits for states, increasing the likelihood that they endorse

TCS demands for adoption. That is, bottom-up pressures should be more effective in fostering top-down support for reform.

This line of reasoning implies the following conditional hypothesis:

H3: The positive relationship between the density of TCS organizations in an institution's issue area and its propensity to adopt MSA reforms is stronger when it performs governance tasks that are costly for states to monitor.

Empirical Analysis
Data and Baseline Specification

The empirical analysis is based on the dataset introduced earlier, an institution-year panel of 2,220 observations (52 institutions covering varying portions of the 1960–2018 period). To capture the dynamic process of reform, the main outcome variable, $\Delta MSA\ Composite$, is the first difference of (i.e., annual change in) *MSA Composite*. I measure issue-specific TCS density in three steps. First, I extract data on the creation and dissolution of INGOs from the Yearbook of International Organizations, the most comprehensive historical catalogue of such organizations.[2] Second, using their listed subject categories and aims, I match INGOs to 20 issue areas delineated by a policy classification scheme proposed by Hooghe et al. (2017). To my knowledge, this represents the first attempt to systematically map the density of INGOs across a large number of domains over time. Third, I use the matched data to construct three treatment variables, which are lagged by 1 year to avoid simultaneity issues: *Issue-Specific INGOs*$_{i,t-1}$, the log number of active INGOs in institution *i*'s issue area(s); *Issue-Specific INGOs, Members*$_{i,t-1}$, the log number of these INGOs that are based in *i*'s member states; and *Issue-Specific INGOs, Largest 5*$_{i,t-1}$, the log number of these INGOs that are based in *i*'s five largest member economies (a number I later vary).

Figure 3 plots the density of INGOs in each issue area between 1960 and 2018. The start of the period is characterized by universally modest density—all areas are the focus of less than 90 INGOs—though there is still some variation across domains (which I later exploit for causal

[2]https://uia.org/yearbook. A small proportion of founding dates and a large proportion of dissolution dates and headquarter locations are missing from the Yearbook. I was able to manually fill in these entries using information from INGO documents and websites, other databases of nongovernmental institutions, and communications with organizational staff.

FIGURE 3 Density of International Nongovernmental Organizations across Issue Areas and over Time
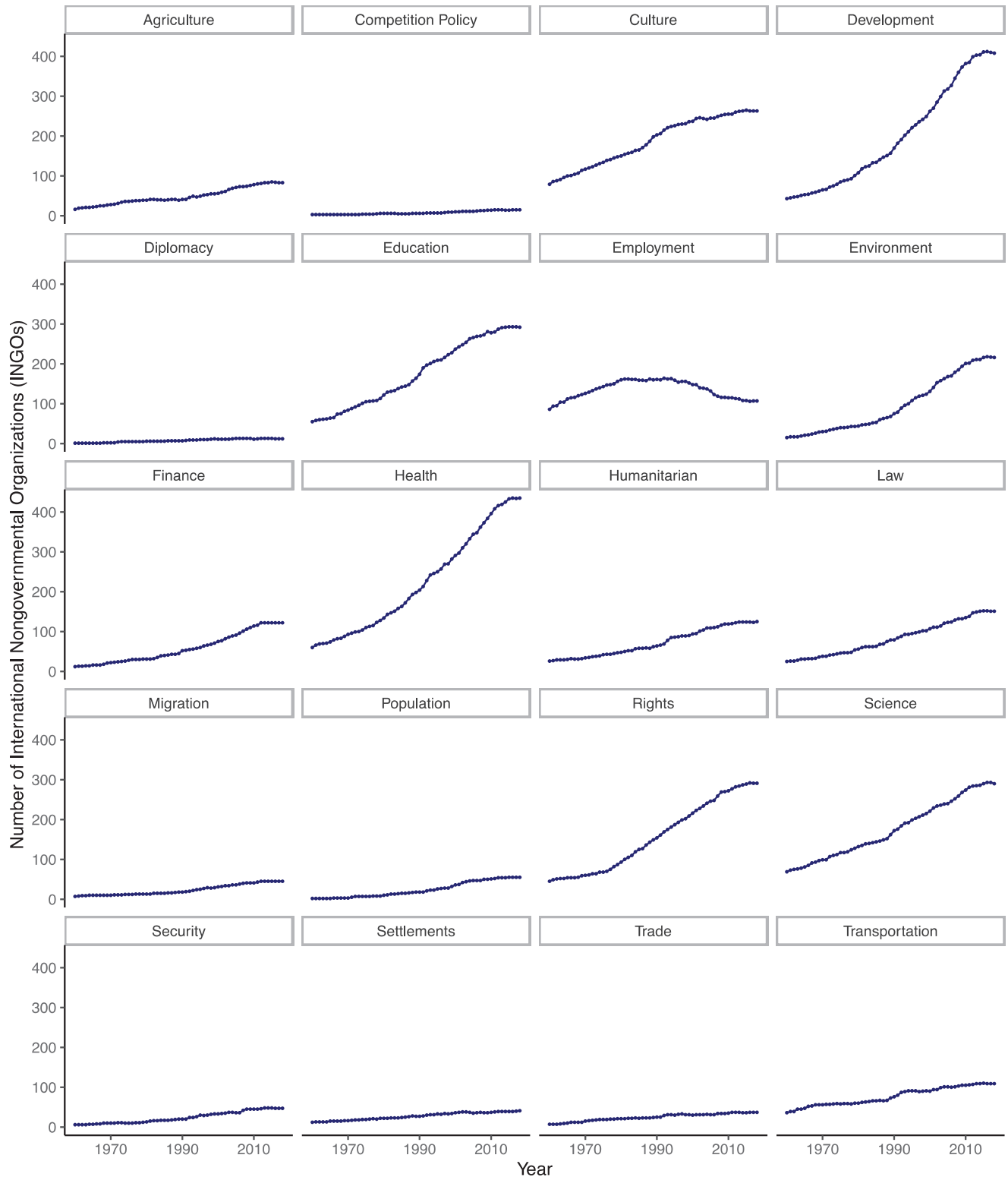
Notes: This figure, which is based on the Yearbook of International Organizations, shows the number of INGOs in 20 major issue areas between 1960 and 2018.

identification purposes). During the subsequent decades, this variation becomes substantially larger with some domains seeing sharp and sustained growth in density and others experiencing negligible change. Growth is highest in the issue areas of economic development and public health, where there are more than 400 INGOs by 2018, and lowest in competition policy and diplomacy, which end the period with less than 20.

I control for the lag of several institution- and time-varying factors analyzed in previous studies of MSA reforms, which can be considered alternative explanations:

- The mean Polity2 score of institution *i*'s member states (*Average Democracy*$_{i,t-1}$) (Marshall, Gurr, and Jaggers 2020). Institutions with a higher proportion of democratic member states may be more willing to adopt MSA mechanisms because accountability is a central principle of democracy (Grigorescu 2010; Tallberg et al. 2013).
- The variance of the foreign policy ideal points of *i*'s member states (*Preference Heterogeneity*$_{i,t-1}$), as estimated by Bailey, Strezhnev, and Voeten's (2017) spatial model of UN General Assembly roll-call votes. Institutions characterized by greater homogeneity in political preferences—and hence fewer "veto players"—may find it easier to reach agreement on accountability reforms (Kahler 2004; Tallberg et al. 2013).
- The average difference between the foreign policy ideal points of *i*'s member states and the United States (*Divergence from US*$_{i,t-1}$), one of the most prominent promoters of oversight reforms in global governance (Grigorescu 2010).
- The log number of *i*'s member states (*Membership Size*$_{i,t-1}$). Institutions with larger memberships, like those with greater preference heterogeneity, contain more veto players (Grigorescu 2010).

My baseline specification, which builds on econometric analyses of the determinants of domestic policy reform, is a panel fixed-effects model estimated with ordinary least squares (OLS) and robust standard errors clustered by institution:

$$\Delta \text{ MSA Composite}_{i,t} = \alpha \text{MSA Composite}_{i,t-1}$$
$$+ \beta \text{INGOs Treatment}_{i,t-1} + \phi \mathbf{X}_{i,t-1} + \gamma_{i,t-1}$$
$$+ \psi_{i,t-1} + \varepsilon_{i,t}, \tag{2}$$

where *INGOs Treatment*$_{i,t-1}$ is *Issue-Specific INGOs*$_{i,t-1}$, *Issue-Specific INGOs, Members*$_{i,t-1}$, or *Issue-Specific INGOs, Largest*$_{i,t-1}$; $\mathbf{X}_{i,t-1}$ is a vector of the control variables described above; γ_i denotes institution fixed effects; and

ψ_t denotes year fixed effects. The fixed effects remove between-institution variation from the analysis, yielding an estimate of the average treatment effect (ATE) *within* institutions over time. They thus help to control for potentially confounding factors that are specific to institutions but tend to vary little over time, such as issue area, mandate, and decision-making rules, as well as specific to years but likely to affect all institutions, such as the global density of TCS, the spread of participatory norms in world politics (Tallberg et al. 2013), and other large-scale international trends. Note, in addition, that I control for the lagged level of *MSA Composite* because the strength of existing MSA mechanisms could affect states' willingness to adopt new ones.

Results

The left column of Figure 4 displays the key baseline estimates. In line with Hypothesis 1, the estimated coefficient on *Issue-Specific INGOs*$_{i,t-1}$ is positive, large, and statistically significant at the 1% level (Model 1). In substantive terms, an approximately 50% increase in the number of INGOs in an institution's issue area is associated with a mean rise in *MSA Composite*$_{i,t}$ of 0.33, which is equivalent to moving from the median value of the outcome variable to the 84th percentile.

The results are similarly strong for the location-specific treatments, although substantive effects are naturally smaller. The coefficients on *Issue-Specific INGOs, Members*$_{i,t-1}$ (Model 2), and *Issue-Specific INGOs, Largest 5*$_{i,t-1}$ (Model 3) are 35% and 48% smaller than that on *Issue-Specific INGOs*$_{i,t-1}$, respectively. This implies that around two-thirds of the ATE in Model 1 is due to INGOs based in member states and more than half is due to INGOs based in the five most powerful of these countries—impressively large proportions.

In the right column of Figure 4, the outcome variable is disaggregated into its five MSA subindices. The results are broadly similar: the coefficients on the three treatments are positive in all 15 models and significant or near significant in 12. The estimates are generally largest and statistically furthest from 0 when Δ *Transparency*$_{i,t}$ and Δ *Redress*$_{i,t}$ are the outcomes. One possible interpretation of this pattern is that proreform groups place greater weight on transparency and redress mechanisms. Another is that institutions are more willing to accede to demands for improvements in these dimensions of accountability than on the evaluation, investigation, and participation fronts.

Full regression estimates for the six sets of models are provided in Appendix C in the online supporting

FIGURE 4 **Relationship between Density of Issue-Specific International Nongovernmental Organizations and Adoption of Multistakeholder Accountability Reforms**
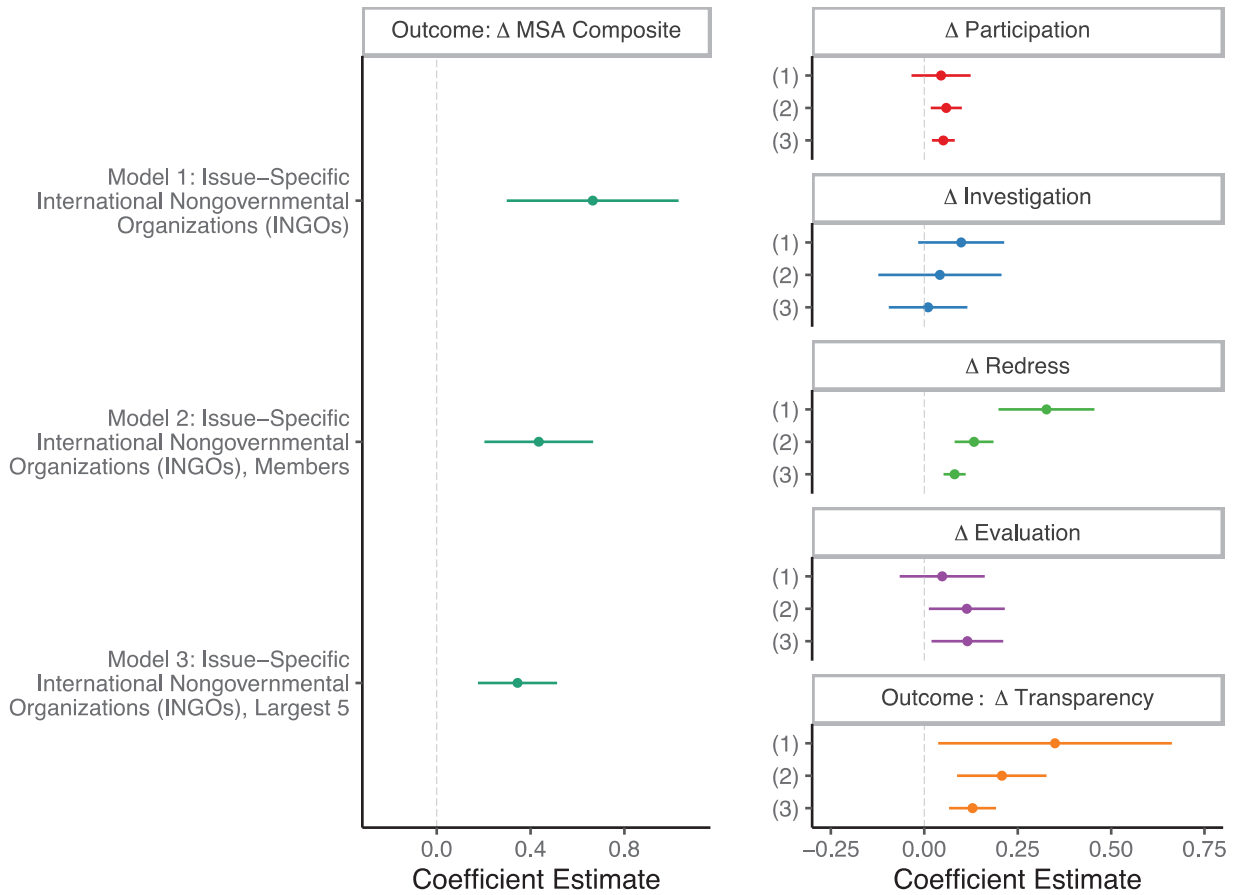
Notes: OLS estimates with 95% confidence intervals based on robust standard errors clustered by institution. Covariates: *Average Democracy$_{i,t-1}$*, *Preference Heterogeneity$_{i,t-1}$*, *Divergence from US$_{i,t-1}$*, *Membership Size$_{i,t-1}$*, *MSA Composite$_{i,t-1}$*, with institution and year fixed effects. $N = 1,952$ in all models.

information. The results for the four control variables are generally weak and inconsistent, with no coefficient attaining significance or maintaining the same sign across a high proportion of models.

Conditional Specification

Turning to Hypothesis 3, I construct dummy variables for whether institutions exercise the six governance tasks enumerated in the previous section from information in annual reports, work programs, and budgets. Over the full sample period, roughly 40% of institutions implement field operations and allocate material resources; one-third design policy interventions, facilitate interstate agreements, and monitor compliance; and four-fifths provide capacity-building assistance. I then modify the

baseline specification by interacting each task dummy with the first treatment variable:

$$
\begin{aligned}
\Delta \text{MSA Composite}_{i,t} &= \alpha \text{MSA Composite}_{i,t-1} \\
&+ \beta \text{Issue-Specific INGOs}_{i,t-1} + \zeta \text{Governance} \\
&\text{Task}_{i,t-1} + \eta \text{Issue-Specific INGOs}_{i,t-1} \\
&\times \text{Governance Task}_{i,t-1} + \phi \mathbf{X}_{i,t-1} + \gamma_i + \psi_t + \varepsilon_{i,t}.
\end{aligned}
$$
(3)

Figure 5 plots the estimated marginal effect of *Issue-Specific INGOs$_{i,t-1}$* on Δ *MSA Composite$_{i,t}$* at both levels of each task dummy. The overall pattern is consistent with Hypothesis 3. The estimates are large and significant at the 95% level for institutions that design interventions, implement operations, and allocate resources—tasks with high monitoring costs for states—but substantially smaller and statistically indistinguishable from 0 for the rest of the sample. In contrast, they are

FIGURE 5 Marginal Effects across Governance Task Moderators
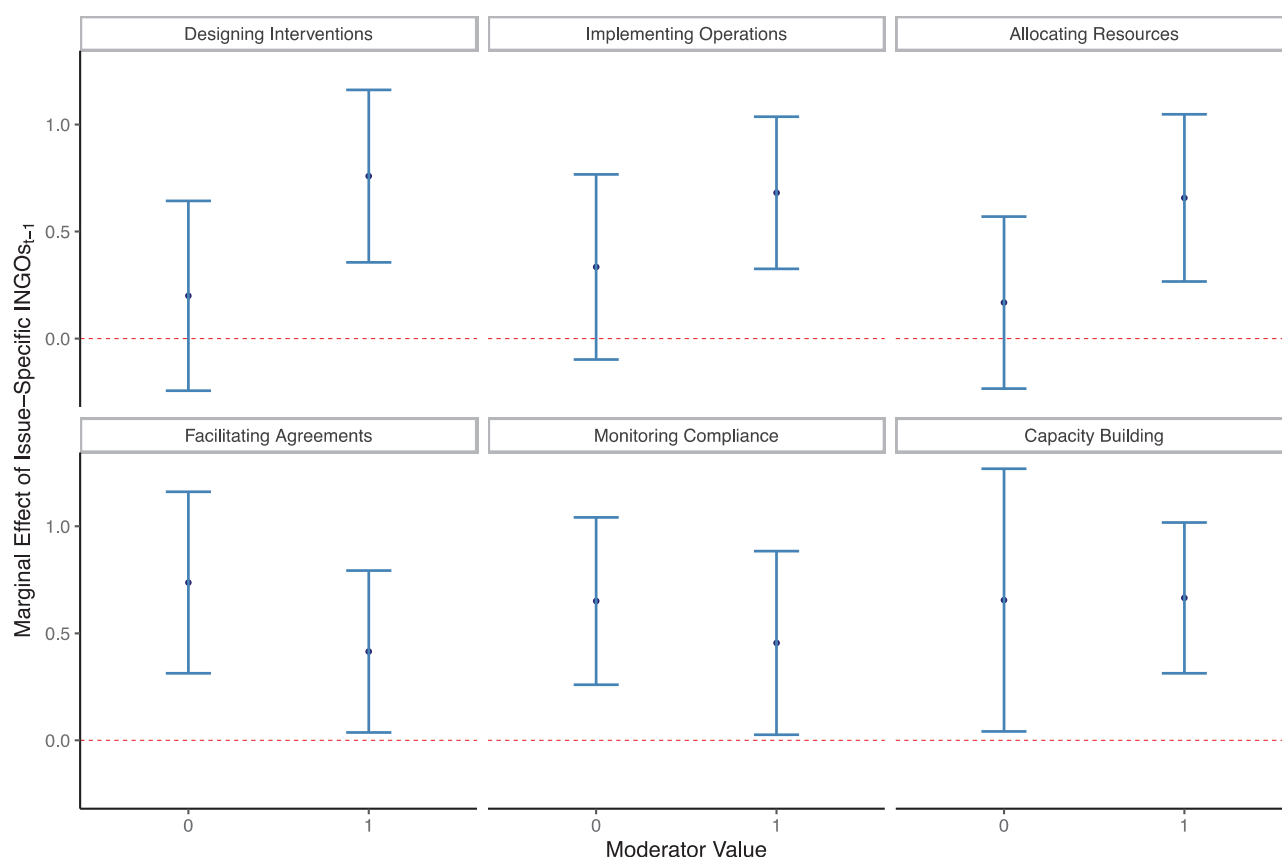
Notes: Estimated marginal effect of *Issue-Specific INGOs$_{i,t-1}$* on Δ *MSA Composite$_{i,t}$* at both levels of each governance task dummy. Estimates are based on the results of Equation (3) (reported in Appendix D in the online supporting information). Vertical bars represent 95% confidence intervals.

smaller when institutions facilitate agreements, monitor compliance, and build capacity—tasks with low monitoring costs—than when they do not, almost losing significance in the case of the first two tasks.

Instrumental Variables Analysis

Although the growth of TCS is generally attributed to advances in communications technology and deepening interdependence between states, it could also be influenced by the opportunities available for engaging with international institutions. It is conceivable, for instance, that the baseline results are driven in part by the formation of INGOs *in response to* ongoing or recent efforts to introduce MSA mechanisms (although not to the overall strength of MSA mechanisms, which is already controlled for). My principal strategy for addressing this possibility is an IV approach that seeks to exploit sources of variation in issue-specific INGO density that are plausibly exogenous

to recent MSA reforms and unlikely to influence current reforms via an alternative causal pathway (thus satisfying the exclusion restriction).

I estimate three IV specifications, the first two of which employ two-stage least-squares (2SLS). In the first specification, the instrument is an interaction between (1) the number of landline and cellular telephone subscriptions per 100 people in the world in year $t-2$, a source of *temporal* treatment variation that captures the global spread of telecommunications technology;[3] and (2) the average annual change in issue-specific INGO density over the three decades prior to the sample (1930–59), a source of *cross-sectional* treatment variation that captures issue-specific determinants of TCS growth. The second instrument, which draws on a common approach to tackling endogeneity in panel settings, is the value of the treatment in year $t-5$. The logic behind this choice is that a lengthy lag of the treatment will not be

[3]Data come from the World Development Indicators database: https://datatopics.worldbank.org/world-development-indicators.

FIGURE 6 Instrumental Variables Estimates
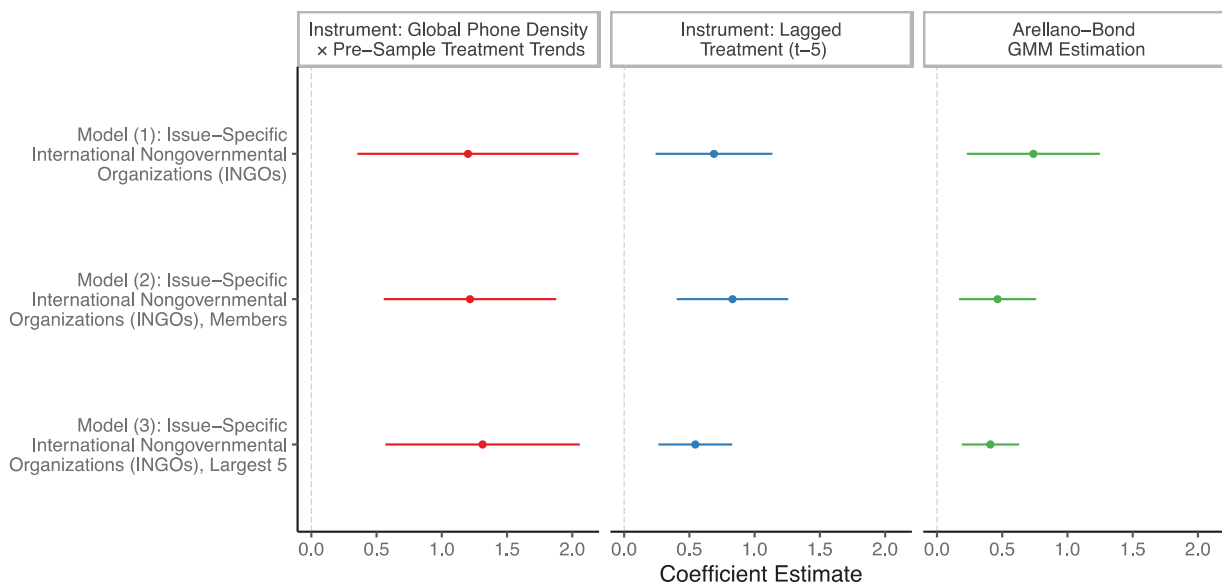
Notes: 2SLS (left and middle panels) and GMM (right panel) estimates with 95% confidence intervals based on robust standard errors clustered by institution. All baseline controls and fixed effects are included in each specification.
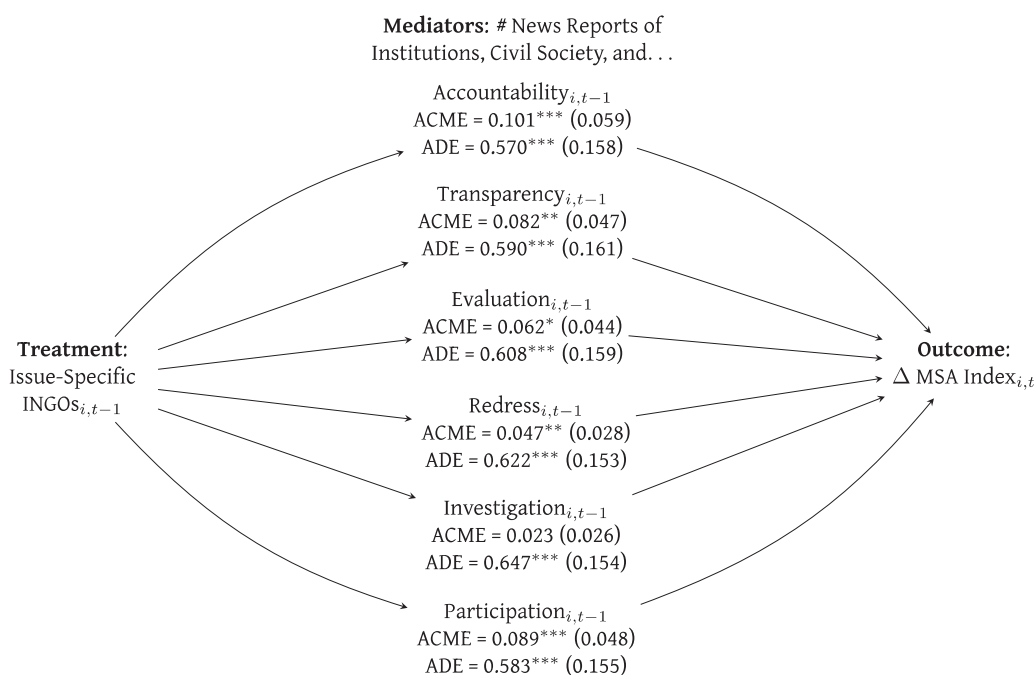
influenced by recent MSA reforms but should nevertheless strongly predict its current value (due to temporal autocorrelation). My third specification employs an Arellano-Bond generalized method of moments (GMM) estimator, another common method for addressing reverse causation that involves including a lag of the outcome as a covariate and instrumenting this variable with progressively longer lags of it. A more extended discussion of each specification and its identifying assumptions is provided in Appendix E in the online supporting information.

As displayed in Figure 6, the three sets of instrumental variables estimates are in line with theoretical expectations. On average, the treatment coefficients are 81% larger than those in the baseline specifications, with the greatest increase in the models featuring the interactive instruments (left column). In other words, the results suggest that any reverse causation in the baseline analysis is likely to have worked *against* rather than for my hypotheses.

Further Robustness Checks

The findings are robust to a number of additional checks, details on which are provided in Appendix F in the online supporting information. First, I add several control variables, some of which again capture previ-

ously suggested alternative explanations: recent MSA reform trends, measured by the mean of the outcome in the 5 years preceding t; the lagged outcome mean among members of the sample (1) based in the same location as i, (2) operating in the same issue area as i, (3) engaged in formal collaborative relations with i, and (4) sharing more than 90% of member states with i, proxies for processes of policy "diffusion" among similar institutions (Grigorescu 2010; Sommerer and Tallberg 2019); the number of IGOs in i's issue area, a measure of institutional competition; and i's income, expenditures, and received contributions. None of the controls are strongly related to the outcome. Second, I convert the baseline specification into an autoregressive distributed lag model (by "undifferencing" the outcome) and an error correction model (by omitting the fixed effects and adding a first-differenced treatment as a regressor). Third, as alternative strategies for dealing with temporal confounding, I (1) control for decade × issue-area fixed effects and (2) replace the year fixed effects with linear and nonlinear time trends. Fourth, I limit (1) the sample to recent decades and to IGOs and (2) the treatment to INGOs focusing on a single issue area. Fifth, to address the possibility of endogeneity in the selection of governance tasks, I instrument the task dummies with their lagged mean among institutions founded near the same time as i, on the intuition that the international political, socioeconomic, and technological factors shaping the most common task choices in each "generation"

FIGURE 7 **Causal Mediation Analysis with Media Coverage as the Mediator**

Mediators: # News Reports of
Institutions, Civil Society, and. . .

$Accountability_{i,t-1}$
ACME = 0.101*** (0.059)
ADE = 0.570*** (0.158)

$Transparency_{i,t-1}$
ACME = 0.082** (0.047)
ADE = 0.590*** (0.161)

$Evaluation_{i,t-1}$
ACME = 0.062* (0.044)
ADE = 0.608*** (0.159)

Treatment:
Issue-Specific
$INGOs_{i,t-1}$

$Redress_{i,t-1}$
ACME = 0.047** (0.028)
ADE = 0.622*** (0.153)

Outcome:
$\Delta\ MSA\ Index_{i,t}$

$Investigation_{i,t-1}$
ACME = 0.023 (0.026)
ADE = 0.647*** (0.154)

$Participation_{i,t-1}$
ACME = 0.089*** (0.048)
ADE = 0.583*** (0.155)

Notes: Average causal-mediation effect (ACME) and average direct effect (ADE) estimates computed using the **mediation** package in R. Cluster-robust standard errors, estimated using a Monte Carlo simulation method, are reported in parentheses.

are credibly exogenous to the more specific dynamics influencing each institution's propensity to adopt—typically decades later—MSA mechanisms. Finally, I conduct a placebo test in which the treatments are replaced by their values for a randomly selected institution in a different issue area, which yields null findings.

Additional Implications

If the argument's causal logic is correct, we should also expect to find a number of more specific empirical patterns that are not captured by Hypotheses 1–3. This section probes three such observable implications.

Media Coverage as a Mediator

First, the argument implies that the positive relationship between the treatment and the outcome will be mediated by evidence of civil society involvement in the MSA reform process. Unsurprisingly, there is no systematic record of such involvement covering all institutions in the dataset. Since generating publicity is a key strategy of transnational advocacy campaigns, however, a reason-

able proxy is the volume of media coverage referencing institutions, civil society groups, and accountability issues. I thus conduct a series of causal-mediation analyses in which the mediators are the log number of (print and online) publications in the Factiva global news database that contain the following three sets of terms:

1. An institution's full name or acronym (unless this acronym is not unique to it);
2. One of the following three strings, in which "?" allows for alternative spellings and "∗" for both singular and plural endings: civil society, nongovernmental organi?ation∗, NGO∗;
3. One of the following six words, which are entered in separate searches to construct six mediators: accountability, transparency, evaluation, redress, investigation, and participation.

Using Imai et al.'s (2011) algorithm for estimating causal-mediation effects, I compute two quantities of interest: the average causal-mediation effect (ACME), the portion of the ATE that is transmitted via the mediator; and the average direct effect (ADE), the remaining portion. The key results, displayed in Figure 7, provide support for a mediation effect. The ACME estimates are positive and significant or close to significant in all six

analyses, accounting for close to one-tenth of the overall ATE, on average. This suggests that MSA-related media coverage is an important—although not dominant—channel through which *Issue-Specific INGOs*$_{i,t-1}$ is related to Δ *MSA Composite*$_{i,t}$.

How Many Powerful States Are Enough?

The theoretical discussion motivating Hypothesis 2b suggested that TCS groups could secure MSA reforms with the support of *varying* numbers of powerful member countries, although noted that case studies of successful campaigns usually identify more than one. To gain a more precise understanding of how many powerful countries are "enough," I now vary the threshold of economic size used to construct the third treatment (*Issue-Specific INGOs, Largest 5*$_{i,t-1}$) from the membership's largest economy to its 10 largest economies. Appendix G in the online supporting information shows that the estimated treatment effect is substantively and statistically similar for thresholds of four or more economies, becoming markedly smaller and almost losing significance below this level. That is, MSA reforms appear to be more likely to occur when TCS groups mobilize pressure on several—rather than a few—powerful member states.

Weighted Voting as a Moderator

Another implication of the discussion around Hypothesis 2b is that TCS pressures on powerful member countries are more effective in inducing reform when these states enjoy greater formal and informal influence in the international policy process. Perhaps the central formal means by which the powerful entrench their control over international institutions is weighted voting rules, which allocate voting shares asymmetrically on the basis of economic and financial criteria. To test whether the density of issue-specific INGOs located in powerful member states is more closely related to MSA reforms when institutions employ such rules, I interact the 10 variants of *Issue-Specific INGOs, Largest*$_{i,t-1}$ described in the previous paragraph with a lagged weighted voting dummy. The results, reported in Appendix G in the online supporting information, are supportive: eight of the 10 interaction terms have positive, large, and highly significant coefficients, and the marginal effects of nine of the 10 treatments are similarly sizable and significant when institutions employ weighted voting but small and nonsignificant when they do not.

Qualitative Evidence

In addition to macrolevel trends in the spread of MSA mechanisms, the argument sheds light on more granular political and strategic processes shaping a number of significant cases and trends described in the second section—processes that cannot easily be captured with quantitative data. This section presents a selection of illustrative examples drawing on institutional reports, case studies, archival records, and 42 interviews with representatives of 14 donor states and officials from 12 institutions.[4]

One of the most conspicuous patterns in the data is the strength of MSA mechanisms in international financial institutions, among which the World Bank has been the earliest and most proactive adopter. The impetus for the World Bank's reforms has consistently come from a broad alliance of TCS organizations and OECD member countries. Indeed, the Narmada scandal only led to the introduction of an access-to-information policy and grievance redress system thanks to a global advocacy campaign orchestrated by well-connected Washington-based INGOs, many of which had been instrumental in the creation of the World Bank-NGO Committee—one of the first mechanisms of civil society consultation in global governance—a decade earlier. Despite stiff opposition from senior management, the campaign successfully pressured the United States, Japan, and Western European nations to use the threat of reduced funding—a threat sharpened by their dominant voting share in the World Bank Executive Board—to secure the institution's consent for transparency and oversight reforms (Fox and Brown 1998).

The receptiveness of OECD policymakers to TCS demands for accountability owed much to the difficulty of overseeing World Bank activities, which encompass all three of the hard-to-monitor governance tasks in my typology. The Narmada fiasco was pivotal to the reform effort precisely because it brought to light how little states knew about the institution's on-the-ground performance. Frustration about this state of affairs was explicitly aired during a congressional hearing on World Bank funding in May 1993, in which politicians and aid officials repeatedly highlighted how their lack of timely and accurate information about projects prevented them from exercising organizational oversight (United States Congress 1993). A similar theme emerges from transcripts of World Bank Executive Board deliberations over TCS proposals for a redress mechanism a few months

---

[4]For further details on the interviews, see Appendix H in the online supporting information.

later (recently obtained, perhaps ironically, via an access-to-information request). In one meeting shortly before the Inspection Panel's creation, for example, the United States' Executive Director emphasized the "important contribution" it would make to "strengthen[ing] the capacity of this Board to conduct its oversight function" (World Bank 1993, 105).

MSA mechanisms inaugurated by the World Bank have been swiftly embraced by other international financial institutions, with intense TCS pressures on OECD governments again playing a decisive role. The battle against the Narmada project was in fact part of a broader transnational advocacy campaign for accountability in MDBs—known officially as the "MDB Campaign"—that began in 1983. Other MDB operations targeted by this campaign include the Asian Development Bank's Theun Hinboun and Nam Theun 2 dam projects in Laos and the Inter-American Development Bank's BR-364 and Polonoroeste road projects in Brazil.[5]

An interesting deviation from this trend is the IMF, which has weaker evaluation and investigation mechanisms than other international financial institutions and, as noted earlier, no grievance redress system. Interviewees from the IMF's secretariat and governing bodies mostly traced these differences to the less intense TCS pressures faced by the institution, which they in turn attributed to its focus on financial and monetary issues (described by one interviewee as "drier" and "less emotive").[6] This explanation is consistent with the data on INGO density presented earlier—historically, far more INGOs have been concerned with development than financial issues—as well as with studies of IMF–civil society relations. Scholte, for instance, notes that "relatively few NGOs have given major priority specifically to the IMF over a sustained period of time" (2002, 13). Nevertheless, interviewees pointed out that even these associations have enjoyed some success in encouraging MSA reforms by exerting pressure on powerful member states. One representative of a large European nation cited the example of the IMF's Independent Evaluation Office, which was created in 2001 after TCS proposals for an evaluation mechanism received the "vital backing" of OECD members of the Executive Board.[7] This support was not difficult to obtain, the delegate added, given

"how hard it is for the Board to monitor and assess financial stabilization programs — from the start of the loan cycle to the end."

Another seemingly puzzling pattern that maps closely onto the distribution of TCS activity across issue areas is the wide variation in MSA structures within the UN System, whose members have similar governance arrangements. UN institutions in domains with a high density of INGOs, such as UNDP and UNEP, have faced stronger TCS pressures for accountability and possess higher *MSA Composite* scores than those in low-density domains, such as the UN Human Settlements Programme (UN-HABITAT) and the UN Office on Drugs and Crime (UNODC). In the words of one senior evaluation officer in UNDP:

> "The UN family isn't a monolithic entity. At one end of the spectrum, we've faced persistent demands from international civil society groups and sympathetic member states to become more transparent, open, and responsive to stakeholders. At the other end, some of our counterparts — especially those working in areas that are of less interest to the general public — have barely faced any such pressures. The end result is significant differences in accountability structures among a fairly similar set of institutions."[8]

Not all variation in MSA structures within the UN System, however, can be explained by issue-specific TCS density alone. For instance, UNCTAD and the UN Industrial Development Organization (UNIDO) have faced comparable bottom-up accountability pressures to other development institutions yet possess far weaker MSA mechanisms, granting almost no access to non-state actors, offering limited avenues for investigation and redress, and lacking any kind of transparency policy. Around the time of the Narmada campaign, both institutions experienced existential crises precipitated by years of policy gridlock and bureaucratic inefficiency, prompting calls for them to improve their accountability and effectiveness or be abolished. In contrast to the World Bank, however, neither institution took meaningful steps to address such demands. Consistent with the argument, member state representatives I interviewed insisted that MSA reforms were not necessary because the institutions' core functions—supporting interstate negotiations and providing technical assistance—were already subject to close scrutiny. One delegate to UNCTAD's governing body, for instance, described its secretariat as performing

---

[5] Author interviews with department director, Asian Development Bank Institute, August 26, 2014, London; and with staff economist, Inter-American Development Bank, April 14, 2014, Washington, DC.

[6] Author interview with member of IMF Independent Evaluation Office, May 14, 2018, Washington, DC.

[7] Author interview with IMF Executive Director, July 14, 2018, Washington, DC.

[8] Author interview with chief of section, UNDP Independent Evaluation Office, May 22, 2018, New York.

"member-facing, member-supporting tasks that we directly observe," rendering "additional oversight channels redundant — despite what some elements of TCS may think."[9]

A similar dynamic has characterized the WTO, whose high-profile contestation by TCS activists yet weak MSA structures make it an intriguing case for the argument. While the WTO may initially appear to defy theoretical expectations, a closer examination reveals that opposition to the institution has occurred in limited and infrequent bursts—the most well-known being the 1999 "Battle of Seattle"—and come mainly from antiglobalization groups concerned with environmental and labor rights issues rather than the substance of trade rules. In reality, relatively few INGOs focus specifically on trade issues (see Figure 3), which may explain why, as one study notes, "demands by NGOs and other outsiders for greater openness [in WTO organs] have been surprisingly sparse" (Stewart and Sanchez Badin 2011, 567). Perhaps even more inimical to reform, however, have been concerns among member states that MSA mechanisms could dilute their control over the policy process without meaningfully enhancing their oversight capacities. These fears emerged clearly during interviews with state representatives, all of whom stressed that MSA reforms were not appropriate for an institution whose chief function is facilitating trade agreements between states. As one delegate to the WTO's General Council rhetorically asked:

> "What would be the point of an Inspection Panel or an Independent Evaluation Office for the WTO? We don't carry out projects in faraway places where staff performance is hard to observe and assess, like the World Bank and the IMF… All the action takes place — and all the staff are based — in Geneva, right in front of our eyes."[10]

Discussion

The emergence of broader and more inclusive mechanisms for holding major international institutions to account is one of the most striking trends in global governance in recent decades. It is not a development, however, that has been embraced by all such institutions. I have argued that a key driving force behind the adoption of MSA mechanisms has been the presence of dense networks of TCS organizations—in particular those located in powerful states—with the capacity to build leverage over member countries through agenda setting, coalition building, and advocacy activities. The effectiveness of these pressures is not constant across institutions, however, but contingent on the costs of monitoring institutions' governance tasks, which affect the strategic gains states enjoy from new avenues for exercising oversight and control.

Original statistical and qualitative evidence has corroborated the argument's main hypotheses as well as a raft of less obvious observable implications. Institutions embedded within denser webs of TCS activity, the results indicate, have opened up more expansive channels for public and private stakeholders to learn about, evaluate, participate in, and impose consequences for their actions, particularly when they perform tasks with high monitoring costs for states. In concrete terms, this could make the difference between consumer groups accessing information and participating in deliberations about a new set of global trade standards, or marginalized communities rectifying and receiving compensation for environmental damage caused by an international development project, and the wholesale exclusion of such actors from institutional processes and operations. As noted earlier, such differences are not only normatively significant in and of themselves but also consequential for other salient characteristics of institutions, such as their legitimacy, performance, and impact.

In emphasizing the interaction between TCS mobilization and strategic state interest, the study complements the few previous examinations of how bottom-up and top-down factors combine to shape global accountability structures. Whereas these analyses focus on variables such as TCS groups' contacts and alignment of interests with domestic policymakers (Pallas and Uhlin 2014) and states' credible commitment problems (Pallas and Urpelainen 2012) and need for nongovernmental resources (Raustiala 1997), my argument draws attention to organizational density as a key determinant of TCS influence and to state concerns about bureaucratic performance as an integral element of the strategic calculus. I find limited empirical support, on the other hand, for a host of purely top-down explanations for cross-institutional differences in the adoption of MSA mechanisms, such as variation in member states' domestic democratic standards, foreign policy preferences, and overall number.

More broadly, the study builds bridges between the influential literatures on transnational advocacy networks, institutional design, and institutional change in international relations, which have spoken to each other surprisingly rarely. First, while access and

[9] Author interview with delegate to UNCTAD Trade and Development Board, June 7, 2012, Geneva.

[10] Author interview with delegate to WTO General Council, June 8, 2012, Geneva.

disclosure regulations have been shown to constrain TCS activities, particularly at the domestic level (Bloodgood 2010), the reverse causal pathway has received relatively little attention. The findings here provide evidence that population-level characteristics of TCS can bring about meaningful change in these as well as other dimensions of formalized accountability in global governance. Second, unlike existing structural theories of institutional change, the argument highlights how the density of organized actors "below"—and not just "beside"—institutions can influence the prospects for governance reforms. Finally, the results show that structural variables can not only impact but also interact in important ways with design features, such as governance tasks and voting rules. In doing so, they point to the potential for productive cross-fertilization between structural approaches, principal-agent analyses of delegation, and power-oriented theories of institutional design.

Appendix: List of International Institutions in Dataset

Full Name	Acronym	Founded
Adaptation Fund	AF	2001
African Development Bank	AFDB	1964
Asian Development Bank	ASDB	1966
Caribbean Development Bank	CDB	1969
Central Emergency Response Fund	CERF	2006
CGIAR	CGIAR	1971
Climate Investment Funds	CIFS	2008
Commonwealth Secretariat	COMSEC	1965
European Bank for Reconstruction and Development	EBRD	1991
European Development Fund	EDF	1959
Expanded Delivering as One Funding Window for the Achievement of the MDGs	EFW	2008

Full Name	Acronym	Founded
Food and Agriculture Organization	FAO	1945
Global Partnership for Education	GPE	2002
Gavi, the Vaccine Alliance	GAVI	2000
Global Crop Diversity Trust	GCDT	2004
Global Environment Facility	GEF	1991
Global Fund to Fight AIDS, Tuberculosis and Malaria	GFATM	2002
Global Facility for Disaster Reduction and Recovery	GFDRR	2006
Inter-American Development Bank	IADB	1958
International Committee of the Red Cross	ICRC	1863
International Fund for Agricultural Development	IFAD	1977
International Finance Corporation	IFC	1956
International Federation of the Red Cross	IFRC	1919
International Labour Organization	ILO	1919
International Monetary Fund	IMF	1945
International Organization for Migration	IOM	1951
International Trade Centre	ITC	1964
Least Developed Countries Fund	LDCF	2001
Multilateral Fund for the Implementation of the Montreal Protocol	MLF	1991
United Nations Office for the Coordination of Humanitarian Affairs	OCHA	1996

Full Name	Acronym	Founded
Office of the United Nations High Commissioner for Human Rights	OHCHR	1994
United Nations Peacebuilding Fund	PBF	2006
Private Infrastructure Development Group	PIDG	2002
United Nations Human Settlements Programme	UN-HABITAT	1978
Joint United Nations Programme on HIV/AIDS	UNAIDS	1994
United Nations Capital Development Fund	UNCDF	1966
United Nations Conference on Trade and Development	UNCTAD	1964
United Nations Development Programme	UNDP	1965
United Nations Environment Programme	UNEP	1972
United Nations Educational, Scientific, and Cultural Organization	UNESCO	1945
United Nations Mine Action Service	UNMAS	1997
United Nations Population Fund	UNFPA	1969
Office of the United Nations High Commissioner for Refugees	UNHCR	1950
United Nations Children's Emergency Fund	UNICEF	1946
United Nations Industrial Development Organization	UNIDO	1966
United Nations Office for Disaster Risk Reduction	UNDRR	2006
UNITAID	UNITAID	2006

Full Name	Acronym	Founded
United Nations Office on Drugs and Crime	UNODC	1966
United Nations Relief and Works Agency for Palestine Refugees in the Near East	UNRWA	1949
UN Women	UNW	1976
World Bank	WB	1944
World Food Programme	WFP	1961
World Health Organization	WHO	1948
World Trade Organization	WTO	1995

References

Abbott, Kenneth W., Jessica F. Green, and Robert O. Keohane. 2016. "Organizational Ecology and Institutional Change in Global Governance." *International Organization* 70(2): 247–77.

Bailey, Michael A., Anton Strezhnev, and Erik Voeten. 2017. "Estimating Dynamic State Preferences from United Nations Voting Data." *Journal of Conflict Resolution* 61(2): 430–56.

Bloodgood, Elizabeth. 2010. "Institutional Environment and the Organization of Advocacy NGOs in the OECD." In *Advocacy Organizations and Collective Action*, eds. Aseem Prakash and Mary K. Gugerty. Cambridge: Cambridge University Press, pp. 91–130.

Buntaine, Mark T. 2015. "Accountability in Global Governance: Civil Society Claims for Environmental Performance at the World Bank." *International Studies Quarterly* 59(1): 99–111.

Ecker-Ehrhardt, Matthias. 2018. "International Organizations 'going public'? An Event History Analysis of Public Communication Reforms 1950–2015." *International Studies Quarterly* 62(4): 723–36.

Fox, Jonathan A., and Lloyd D. Brown, eds. 1998. *The Struggle for Accountability: The World Bank, NGOs, and Grassroots Movements.* Cambridge: MIT Press.

Grant, Ruth W., and Robert O. Keohane. 2005. "Accountability and Abuses of Power in World Politics." *American Political Science Review* 99(1): 29–43.

Grigorescu, Alexandru. 2007. "Transparency of Intergovernmental Organizations: The Roles of Member States, International Bureaucracies and Nongovernmental Organizations." *International Studies Quarterly* 51(3): 625–48.

Grigorescu, Alexandru. 2010. "The Spread of Bureaucratic Oversight Mechanisms across Intergovernmental Organizations." *International Studies Quarterly* 54(3): 871–86.

Hawkins, Darren G., David A. Lake, Daniel L. Nielson, and Michael J. Tierney, eds. 2006. *Delegation and Agency in International Organizations.* Cambridge: Cambridge University Press.

Honig, Dan, Ranjit Lall, and Bradley C. Parks. 2022. "When Does Transparency Improve Institutional Performance? Evidence from 20,000 Aid Projects in 183 Countries." *American Journal of Political Science* https://doi.org/10.1111/ajps.12698.

Hooghe, Liesbet, Gary Marks, Tobias Lenz, Jeanine Bezuijen, Besir Ceka, and Svet Derderyan. 2017. *Measuring International Authority: A Postfunctionalist Theory of Governance, Volume III.* Oxford: Oxford University Press.

Imai, Kosuke, Luke Keele, Dustin Tingley, and Teppei Yamamoto. 2011. "Unpacking the Black Box of Causality: Learning about Causal Mechanisms from Experimental and Observational Studies." *American Political Science Review* 105(4): 765–89.

Jupille, Joseph, Walter Mattli, and Duncan Snidal. 2013. *Institutional Choice and Global Commerce.* Cambridge: Cambridge University Press.

Kahler, Miles. 2004. "Defining Accountability Up: The Global Economic Multilaterals." *Government and Opposition* 39(2): 132–58.

Keck, Margaret E., and Kathryn Sikkink. 1998. *Activists beyond Borders: Advocacy Networks in International Politics.* Ithaca: Cornell University Press.

Kingsbury, Benedict, Nico Krisch, and Richard B. Stewart. 2005. "The Emergence of Global Administrative Law." *Law and Contemporary Problems* 68(3/4): 15–61.

Lall, Ranjit. 2017. "Beyond Institutional Design: Explaining the Performance of International Organizations." *International Organization* 71(2): 245–80.

Lall, Ranjit. 2023. *Making International Institutions Work: The Politics of Performance.* Cambridge: Cambridge University Press.

Marshall, Monty, Ted Robert Gurr, and Keith Jaggers. 2020. "*Polity V Project: Political Regime Characteristics and Transitions, 1800–2018.*" Vienna, VA: Center for Systemic Peace.

Martin, Lisa. 2006. "Distribution, Information, and Delegation to International Organizations: The Case of IMF Conditionality." In *Delegation and Agency in International Organizations,* eds. Darren G. Hawkins, David A. Lake, Daniel L. Nielson, and Michael J. Tierney. Cambridge: Cambridge University Press, pp. 140–64.

Nielson, Daniel L., and Michael J. Tierney. 2003. "Delegation to International Organizations: Agency Theory and World Bank Environmental Reform." *International Organization* 57(2): 241–76.

Pallas, Christopher L., and Anders Uhlin. 2014. "Civil Society Influence on International Organizations: Theorizing the State Channel." *Journal of Civil Society* 10(2): 184–203.

Pallas, Christopher L., and Johannes Urpelainen. 2012. "NGO Monitoring and the Legitimacy of International Cooperation: A Strategic Analysis." *The Review of International Organizations* 7(1): 1–32.

Raustiala, Kal. 1997. "States, NGOs, and International Environmental Institutions." *International Studies Quarterly* 41(4): 719–40.

Scholte, Jan Aart. 2002. "*Civil Society Voices and the International Monetary Fund.*" Ottawa: North-South Institute.

Scholte, Jan Aart, ed. 2011. "*Building Global Democracy?*" *Civil Society and Accountable Global Governance.* Cambridge: Cambridge University Press.

Sommerer, Thomas, and Jonas Tallberg. 2019. "Diffusion Across International Organizations: Connectivity and Convergence." *International Organization* 73(2): 399–433.

Stewart, Richard B., and Michelle R. Sanchez Badin. 2011. "The World Trade Organization: Multiple dimensions of Global Administrative Law." *International Journal of Constitutional Law* 9(3-4): 556–86.

Tallberg, Jonas, Thomas Sommerer, Theresa Squatrito, and Christer Jönsson. 2013. *The Opening Up of International Organizations.* Cambridge: Cambridge University Press.

United States Congress. 1993. "Authorizing Contributions to IDA, GEF, and ADF." *Subcommittee on International Development, Finance, Trade, and Monetary Policy.* Washington, DC: US Congress.

Woods, Ngaire. 2001. "Making the IMF and the World Bank more accountable." *International Affairs* 77(1): 83–100.

World Bank. 1993. "Meeting of the Executive Directors." Transcript, July 9.

Supporting Information

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

Appendix A: Description of Data Sources
Appendix B: Additional Information on Dataset
Appendix C: Baseline Results
Appendix D: Conditional Results
Appendix E: Instrumental Variables Analysis
Appendix F: Further Robustness Checks
Appendix G: Testing Additional Implications
Appendix H: Interview Evidence