Supplementary Information for the article "The International Authority Database"

Contents

A Measuring IO Authority **2**
- A.1 Authority Concept and Operationalization 2
- A.2 Coding rule-making authority: FAO 7
- A.3 Aggregation 10

B Sample **12**
- B.1 List of IOs 12
- B.2 Sampling 14

C Variables in Dataset **20**

D Contrast to MIA **22**

A Measuring IO Authority

This section provides technical supplementary information regarding the coding and aggregation process. We start with a general operationalization scheme of authority as defined in the main text. Next, we show how we adapted this general scheme to code specific policy functions. We illustrate with the routine compliance monitoring and rule-making functions. Then, we present in detail how we coded the FAO's rule-making authority, discussing also numerical codes and aggregation procedures. The section concludes with an illustration of how we aggregated all information into a single authority score.

A.1 Authority Concept and Operationalization

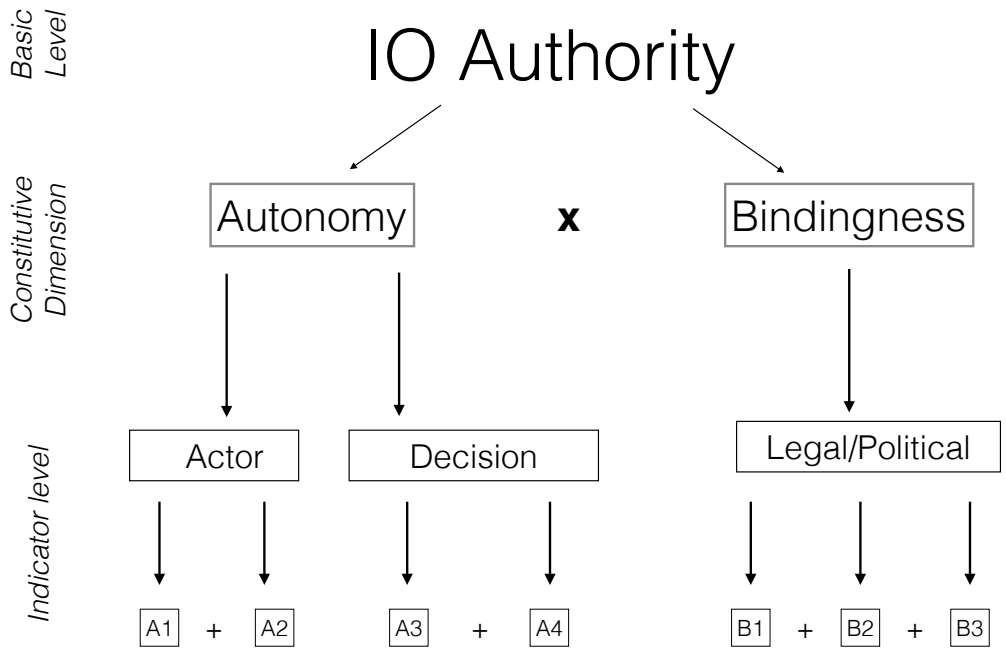

Figure SI.1: Authority Operationalization

Figure SI.1 presents a graphical scheme of our conceptualization and operationalization of IO authority. The graphical presentation follows Goertz (2006). Authority, our basic-level con-

cept that we aim to measure, is constituted by two dimensions: Autonomy and Bindingness. These are multiplicatively related, indicated by the **x**. Both dimensions, in turn, are characterized by different institutional indicators. For example, the type of Actor and the Decision rules indicate an IO's autonomy.

Whether and to what extent rules are legally binding or decisions politically binding serves as an indicator for the Bindingness dimension. Each of these indicators can have different values, as demonstrated by the small rectangles on the bottom of Figure SI.1. For instance, rectangle 'A1' could be an *assembly* of states and rectangle 'A3' a *unanimity* voting rule. Rather than suggesting addition, the + signs between the lowest-level rectangles indicate a logical 'or' (Goertz, 2006).

This general scheme is applied to all seven policy functions. However, for each function, we adapt the indicator level to match the relevant institutional characteristics and procedures of each function. As already noted in the main text, each function differs in terms of autonomy and bindingness indicators.

We schematically present the operationalization of rule-making and monitoring authority in Figure SI.2. As noted, we measure the autonomy of Rule Making by determining the IO body that adopts policy instruments and its corresponding voting rule (top panel in Figure SI.2). The values on these two indicators are combined according to the procedure laid out in the main text (see Table SI.4 on page 9). Bindingness of rule making is indicated by whether the adopted policy instrument is legally binding for states *after* adoption. One possible value is conditionally binding (e.g. ratification required).

The bottom panel in Figure SI.2 adapts the general scheme to the routine compliance monitoring function. In total, we differentiate between four types of monitoring. First, we distinguish between compliance monitoring (respecting state obligations) and performance monitoring (following recommendations). Second, within each of these two monitoring objectives, we further differentiate between routine monitoring (on a regular basis) and special investigations (for serious violations in the case of compliance monitoring or for special recommended measures in case of performance assessment). This results in four different types of monitoring. Here, we focus on routine compliance monitoring.

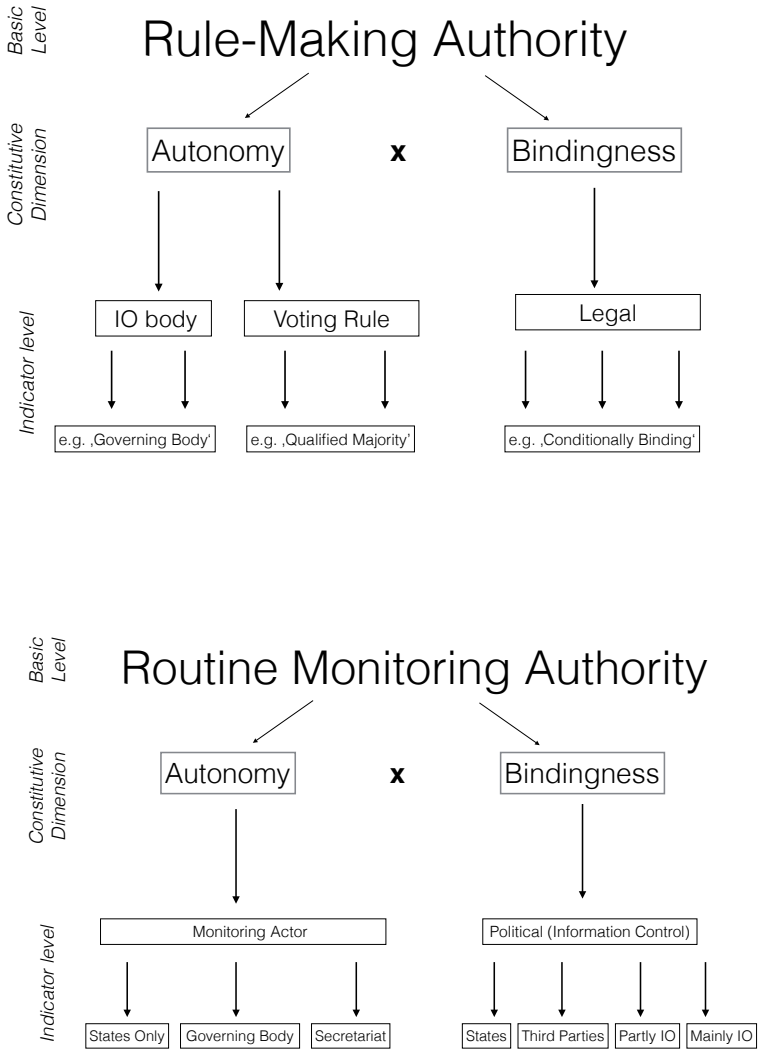

Figure SI.2: Operationalization of Rule-Making (top panel) and Routine Compliance Monitoring (bottom panel) functions

In contrast to rule making, no extra decision by an IO organ is required to carry out routine compliance monitoring. Routine compliance monitoring forms part of the core tasks states transferred to IOs via constitutive legal documents (if at all). Therefore, the indicators for IO autonomy concerning routine compliance monitoring differ from the ones in rule making. IO monitoring autonomy is indicated by the type of actor that is entitled to carry out compliance monitoring. As such, the bottom panel of Figure SI.2 presents different values for the indicator on 'Monitoring Actor'.

Similarly, the indicator for bindingness differs from the one in rule making. Instead of asking whether routing compliance monitoring is *legally* binding—it is, by definition—we assess the extent to which routine compliance monitoring provisions are politically binding in terms of costs they impose upon states. The costs are a function of a member state's (lack of) control over compliance-relevant information. Therefore, we ask about information sources and instruments available to IOs in order to assess state parties' compliance with substantive rules. The less each state controls compliance-relevant information, the more binding monitoring is.

In terms of operationalization, we proceed analogously with all remaining functions. Having defined each function's autonomy and bindingness indicators, we then proceed to develop coding schemes. These coding schemes assign ordinal values/codes to each indicator and allow us to systematize raw information from IOs' legal texts.

We illustrate our coding in the following paragraphs. Table SI.1 presents the coding scheme for routine compliance monitoring. The first column of Table SI.1 contains the constitutive dimensions of authority, autonomy, and bindingness. For each dimension, we define the relevant indicator(s). Each indicator can have one of several codes or values. These codes/values are ordered from low to high in terms of the relevant dimension. Finally, the numbers in the column 'Values' represent this ordering.

More specifically, we first code whether legal texts provide any procedures or rules regarding routine compliance monitoring. If it is not, the entire function receives a value of zero (i.e. the code 'Not in agreement'). Otherwise, we code the type of 'Monitoring actor' to obtain information on the autonomy dimension of routine compliance monitoring. The assembly of states may carry out monitoring or, in very few instances, it may create a subsidiary body on

its behalf to conduct monitoring. If it does, we record it by the code 'States only'. We assign the code 'Governing body involved' if the IO's governing body is tasked to monitor compliance, also jointly with permanent IO staff. Finally, we assign the code 'Only Secretariat' if the IO secretariat is exclusively responsible for monitoring compliance on a routine basis. This expresses the highest degree of autonomy.

Table SI.1: Coding scheme for monitoring function (routine compliance monitoring)

Dimension	**Indicator**	**Codes**	**Values**
Autonomy	Monitoring actor	Not in agreement	0
		States only/Assembly	1
		Governing body involved	2
		Only Secretariat	3
Bindingness	Information sources and instruments	From states only	0
		Also from third parties	1
		Partly IO	2
		Mainly IO	3

Regarding the bindingness of routine compliance monitoring, we deviate from a legal evaluation of the monitoring provisions. By definition, the substantive provisions of IOs and the monitoring procedures are legally binding upon states, as they are formalized in the legally binding founding treaties of IOs. Adapting the bindingness dimension to compliance monitoring requires us to assess the extent to which the process of compliance monitoring limits state discretion. To that end, we focus on the information and monitoring instruments that an IO is entitled to use when it evaluates member state compliance. The less an IO depends on states' input of compliance-relevant information, the higher the bindingness of monitoring.

The first code we assign captures whether an IO can only use reports submitted to it by states to assess their compliance. In these cases, the IO mostly functions as a distributor of state-generated information, exchanging country reports among its membership. This barely limits state discretion of what and how accurately to report. We assign a second code when an IO can, in addition to state-provided information, use available reports by NGOs or other IOs ('Also from third parties').

In contrast to the first two codes, the third one captures instances where an IO has its own sources of information to monitor compliance. This applies, for example, when the IO can request state parties to submit additional reports. Still, the primary sources of its compliance information come from states, NGOs, or other third parties.

Finally, we assign the fourth code ('Mainly own') if an IO has its own, independent toolbox of monitoring instruments to generate compliance information, possibly supplemented by additional sources of information (e.g. NGOs, other IOs). Having the competence to conduct on-site inspections or to continuously oversee certain member state activities are two possibilities of how IOs can generate their own compliance information.

This section has shown how we operationalize our authority concept by defining relevant indicators for each of the seven policy functions. Moreover, we have presented how we translate the conceptualization into a coding scheme that allows us to identify and systematize relevant information from IOs' legal texts. Finally, the section has also demonstrated that we do not apply the same coding scheme to all functions, but rather that we tailor our general scheme to each function individually in order to obtain valid indicators of autonomy and bindingness.

A.2 Coding rule-making authority: FAO

In this section, we provide technical details on our coding process, continuing with the example of how we coded the FAO's rule-making authority from the main text. We not only show how we assigned codes to particular institutional features, but also how we aggregate them. The section concludes with an illustration of how we arrived from function-level scores per IO to an overall IO authority value.

We begin coding the FAO's rule-making authority with the coding scheme as shown in Table SI.2.

As noted in the main text, we start coding by checking whether an IO's foundational legal texts contain any rule-making provisions at all. If it does not, we assign the code 'Not in agreement'. Otherwise, we identify all policy instruments that an IO can adopt. Having identified and listed all such instruments, we begin by coding an IO's rule-making autonomy. The first step consists of determining the IO body that adopts the respective policy instrument. In the

Table SI.2: Coding scheme for rule making function

Dimension	Indicator	Codes	Values
Autonomy	IO body type	Not in agreement	0
		Assembly	1
		Governing body	2
		Independent organ involved (e.g. secretariat)	3
	Voting rule	Not in agreement	0
		Unanimity	1
		Qualified Majority	2
		Simple Majority	3
Bindingness	Legal bindingness of policy instrument	Not binding	0
		Conditional	1
		Directly binding	2

case of the FAO, the assembly of states adopts *Conventions*, while the FAO's Executive Commission, i.e. the FAO's governing body, adopts *Supplementary Agreements*. Along with the IO body type, we code the corresponding voting rule for adopting each specific instrument.[1] Finally, we assess whether the policy instrument is 'Not binding', 'conditionally binding', or 'directly binding'.

Having identified the relevant information, we then proceed to systematize that information on rule-making procedures. Concerning the FAO, Table SI.3 shows the codes and values for each policy instrument.[2]

Table SI.3: Raw codes: Autonomy and Bindingness of FAO Rulemaking. Numbers in parentheses indicate numerical codes in Table SI.2

IO	Policy Instrument	IO body	Voting Rule	Legal bindingness
FAO	Conventions	Assembly (1)	Qualified majority (2)	Conditional (1)
FAO	Supp. Agreements	Governing body (2)	Qualified majority (2)	Conditional (1)

With that systematized raw information, we then proceed to summarize information at the level of constitutive dimensions. In the case of rule making, we combine information only on

[1] If no voting rule is defined for a given instrument, we code the default voting rule of the respective IO body.
[2] The FAO has more policy instruments than the two indicated. Showing only two serves illustrative purposes.

'IO body' and 'Voting Rule' to arrive at a single number measuring the autonomy dimension of rule making. We do so, because rule-making autonomy is composed of two distinct indicators, while only one indicator measures bindingness (extent of legal bindingness).

Table SI.4 shows how we aggregate information on IO body type and voting rule into a single autonomy score.

Table SI.4: Cross-tabulation and rank ordering of IO body type and voting rule

	Unanimity	Qualified Majority	Simple Majority
Assembly	1	2	3
Governing Body	4	5	6
Independent Body	7	8	9

The combination of 'Assembly' and 'Qualified Majority' for FAO *Conventions* produces an aggregate autonomy score of 2. In turn, 'Governing Body' and 'Qualified Majority' for FAO *Supplementary Agreements* results in an autonomy score of 5. Because both policy instruments are conditionally binding, they both receive a bindingness score of 1, according to Table SI.2 and SI.3.

Columns 'Autonomy' and 'Bindingness' in Table SI.5 display the final dimension scores for each policy instrument of the FAO.

Table SI.5: Obtaining authority scores: Rule making in FAO in the year 2013

IO	Policy Instrument	Autonomy	Bindingness	Authority	Normalized
FAO	Conventions	2	1	$\sqrt{2 \times 1} = 1.41$	0.33
FAO	Supp. Agreements	5	1	$\sqrt{5 \times 1} = 2.23$	**0.53**

As said in the main text and shown in the general operationalization scheme of authority in Figure SI.1, we use the geometric mean to aggregate autonomy and bindingness scores. More formally,

$$\text{Authority}_{ijt} = \sqrt{\text{Autonomy}_{ijt} \times \text{Bindingness}_{ijt}}$$

where i is the IO, j the policy function with $j = 1, \ldots, 7$, and t the year. Aggregating our data with the geometric mean has two key advantages; one conceptual and one methodological. First, we incorporate the conceptual non-substitutability of autonomy and bindingness into our

measure. That is, both need to be present and positive-valued to constitute authority (this is achieved by taking the product of autonomy and bindingness). Second, the geometric mean allows us to produce a single score by combining differently measured variables. As the graphical conceptualizations show, autonomy and bindingness have been measured using a varying number of indicators. Therefore, the resulting value ranges differ as well. The geometric mean allows us to combine them by retaining their original measurement scale and still produce a consistent single score. As a result, a ten percent increase in bindingness, for example, translates into a ten percent increase of the final authority measure. Likewise, a similar increase in autonomy translates into a corresponding increase in authority. The same does not apply when calculating the arithmetic mean of differently measured variables. Formally, the geometric mean is the *n-th* root of the product of *n* variables. As we multiply two variables, we thus take their *square* root.

The column 'Authority' in Table SI.5 exemplifies the calculation of the authority score. The last column normalizes the authority score to range between 0 and 1.[3] The procedure laid out in Table SI.5 gives an authority score per FAO policy instrument. Note that this does not represent the final function-level authority score (Authority_{ijt}). To arrive at the final authority score, we pick the instrument with the highest score, which is *Supplementary Agreements* in the case of the FAO in 2013 (normalized authority of .53). We do so because we want to get at the maximum extent of authority that states transfer to IOs.

A.3 Aggregation

We proceed analogously in coding and aggregating information for all other functions. As a result, we obtain an authority score for each IO, year, and function. Table SI.6 shows the function-level authority scores of the FAO for the year 2013. To arrive at the final IO-level aggregate authority score, we simply sum the individual scores of each of the seven policy functions.

More formally,

[3] Whether we normalize the individual dimensions or the geometric mean is mathematically equivalent.

$$\text{Authority}_{it} = \sum_{j=1}^{7} \sqrt{\text{Autonomy}_{ijt} \times \text{Bindingness}_{ijt}}$$

We thus arrive at an aggregate *single* authority score per IO (i) and year (t). The resulting sum is normalized to range between 0 and 1. The maximum score that an IO can possibly attain is 1, the minimum score being 0.

Table SI.6: Aggregating function-level authority scores: FAO in 2013

IO	Policy function	Normalized authority score
FAO	Agenda setting	.77
FAO	Rule making	**.53**
FAO	Monitoring	.28
FAO	Knowledge generation	.82
FAO	Norm interpretation	.36
FAO	Enforcement	.00
FAO	Evaluation	.70
Sum		3.46
Normalized total authority		.497

B Sample

This section presents our sample of IOs and the procedure to arrive at a geographically and thematically representative sample of IOs.

B.1 List of IOs

Table SI.7 presents the composition of our sample. The start dates refer to the year when the respective IO started its operation—not the year when the treaties creating the IOs were adopted. Most of the abbreviations are official acronyms, with some exceptions (e.g. Commonwealth Secretariat and Nordic Council).

Table SI.7: IAD IO Sample

Name	Abbreviation	Start Year	End Year
Association of Southeast Asian Nations	ASEAN	1967	2013
African Union	AU	1963	2013
Andean Community	Andean	1969	2013
Bank for International Settlement	BIS	1930	2013
Central African Economic & Monetary Union	CEMAC	1994	2013
Caribbean Community	CARICOM	1973	2013
Council of Europe	CoE	1949	2013
Commonwealth Secretariat	Commsec	1965	2013
European Union	EU	1958	2013
Food and Agriculture Organization	FAO	1945	2013
International Criminal Court	ICC	1998	2013
International Coffee Organization	ICO	1963	2013
Intergovernmental Authority on Development	IGAD	1996	2013
International Labour Organization	ILO	1919	2013
International Monetary Fund	IMF	1944	2013
International Whaling Commission	IWC	1948	2013
North Atlantic Fisheries Organization	NAFO	1979	2013
North American Free Trade Agreement	NAFTA	1994	2013
North Atlantic Treaty Organization	NATO	1949	2013
Nordic Council	NC	1962	2013
Organization of Arab Petroleum Exporting Countries	OAPEC	1968	2013
Organization of American States	OAS	1951	2013
Organization for Economic Cooperation and Development	OECD	1961	2013
Organization of Islamic Conference	OIC	1972	2013
Organization for Security and Co-operation in Europe	OSCE	1975	2013
Pacific Island Forum	PIF	1973	2013
Southern African Development Community	SADC	1993	2013
Shanghai Cooperation Organization	SCO	2003	2013
Arab Maghreb Union	UMA	1989	2013
UN Education, Scientific, and Cultural Organization	UNESCO	1946	2013
United Nations	UN	1945	2013
International Bank for Reconstruction and Development	WB	1945	2013
World Health Organization	WHO	1948	2013
World Trade Organization	WTO	1994	2013

B.2 Sampling

Our goal is to construct a cross-section time-series dataset that informs us about the distribution of authority across IOs and across each of the seven policy functions within each IO over time. Time is recorded annually, starting with the date of effective IO operation and ending in the year 2013. Determining autonomy and bindingness per policy function and IO requires a comprehensive and detailed assessment of institutional designs (see above).

For this reason, our coding instrument features more than 100 items that code formal rules and procedures across several organizational bodies (e.g. assemblies, governing bodies, consultative bodies, courts) and policies of an IO (e.g. for the EU these would be, among others, common market policies or decisions regarding the Common Foreign and Security Policy). To code these items, we use the constitutive treaties of IOs but go well beyond these by considering a range of other relevant legal documents: rules of procedure of IO bodies, statutes, protocols, codified political declarations, and annexes to conventions. Because the number and content of legal documents varies over time, we code their amendments and new legal documents as well.

Collecting this information on the exercise of IO authority requires an in-depth coding of a large number of legal documents. Necessarily, then, we sample a fraction of IOs from the relevant population to perform our in-depth coding and to ensure the right balance between sample size and measurement validity. During the sampling process, we pay particular attention to the following two aspects: defining the relevant population and drawing a representative sample from it.

Notwithstanding the many options to obtain a list of IOs (e.g. Yearbook of International Organizations) and use it as a relevant population, we use the population list composed by Sommerer and Tallberg (2017). This population includes IOs that feature the necessary conditions for exercising authority as we define it. Therefore, it is close to our purposes. The criteria for inclusion in the population have been laid out in the main text. Applying these four definitional criteria, Sommerer and Tallberg (2017) considerably reduce the number of relevant IOs in the population. The activity criterion, for example, eliminates all 'dead' (e.g. League of Nations) or dormant IOs that might exist on paper but that have not convened in the last years or that have not produced any output. Being intergovernmental and independent, in turn, excludes

not only emanations (IOs created by other IOs), but also those seemingly autonomous IOs that are in fact large sub-bodies of other IOs (e.g. UNICEF, European Central Bank). Finally, counting IOs with predecessors as one organization (e.g. Organization for African Unity, AU) further reduces the relevant population of IOs. The population list of Sommerer and Tallberg (2017) contains 182 IOs. Ours counts 174 IOs. This difference is due to two reasons: First, Sommerer and Tallberg (2017) double-count certain cross-regional IOs (7) by assigning them to two regions simultaneously. We assign each IO to only one region (see below). Second, we update the IO activity criterion to 2012 with the result that the Western European Union drops from the population list (dissolved 2011).

Our relevant population of IOs counts 174 organizations, comprising also commodity organizations and all regional economic bodies that conform to the definitional criteria. In order to ensure a representative sample, we categorize each IO in the population into one out of 20 different strata. The stratification helps to minimize selection biases. A long tradition of institutional scholarship has provided compelling evidence that issue area and regional dynamics affect institutional outcomes. Hence, organizations will also be differently distributed across world regions and thematic issue areas. As a result, these two variables define our strata (see also Sommerer and Tallberg, 2017).

Following the UN classification of world regions, we define five regions: Africa, Americas (North, Central, South), Asia & Oceania, Europe, and the World. Unlike the UN, we extend the Asian region by Oceania and thus obtain the World as the fifth region, encompassing all IOs with a global scope and mandate. There are cases where an IO's mandate can span over two regions (e.g. the SCO with East European and Asian member states). In such cases, we assign the IO to the region where most of its members are located, given that the organization is not active beyond its region (e.g. NATO). The SCO would thus count as an Asian organization.

We define four broad issue areas. The first issue area comprises all IOs that foster and regulate markets (goods and services) or economic transactions more generally between states and societies. We label this issue area 'Economy'. The World Bank, the International Sugar Union, or the International Telecommunication Union figure in that issue area. Organizations that protect and promote human lives, livelihoods, and liberties, including a clean environment

and cultural self-determination, form the second issue area, which we briefly label 'Human Rights'. Apart from organizations that directly promote human rights (ICC), we also include environmental organizations (e.g. Global Environmental Facility), and those that advance cultural exchanges and identities (e.g. Union Latina). Security organizations demarcate the third issue area and comprise alliances, disarmament and nonproliferation IOs (e.g. International Atomic Energy Agency), and human security organizations (e.g. International Criminal Police Organization). Finally, we group all IOs that engage in at least two distinct issue areas into the multi-issue category.

Combining five regions and four issue areas results into 20 distinct strata and gives us a proportional allocation of population IOs across these strata. For example, 11.5% of all IOs in the population are African economic IOs, such as the East African Development Bank or the African Timber Organization. These population shares give us precise sampling targets and can be used to construct sampling adjustment weights to further balance the sample.

We sample 20% of the population IOs through stratified random sampling using a random numbers generator in standard statistical software (STATA) and sample without replacement based on the strata-specific distribution. This gives us 34 IOs in total, representative of the regional and issue area distribution in the relevant population and of sufficient size to conduct a comprehensive and in-depth analysis of authority patterns, both across IOs and their functions, as well as over time.

Table SI.8: IO distribution in population (P) and sample (S). Cell entries denote percentages. The size of our sample (n) is 34 IOs. The population size (N) is 174 IOs.

	Economy		**Human Rights**		**Security**		**Multi-issue**		***Total***	
	P	S	P	S	P	S	P	S	P	S
Africa	11.5	11.7	4	3	0	0	2.3	2.9	17.8	17.6
America	10.3	9	2.3	3	1.1	0	2.3	2.9	16	14.9
Asia	6.3	6	1.1	0	1.1	3	2.3	2.9	10.8	11.9
Europe	7	6	6	6	0.5	0	2.8	2.9	16.3	14.9
World	23	23	11	11.7	3.4	3	1.7	2.9	39.1	40.6
Total	58.1	55.7	24.4	23.7	6.1	6	11.4	14.5	100	100

To analyze the alignment between the population and our sample, we presented in the main text the absolute number of IOs to sample and actually sampled from each stratum. Given

that distribution, our sample is representative of the relevant population of IOs. Table SI.8 shows the population and sample *shares* of IOs across strata. As can be seen, the relative distributions align rather well. The multi-issue stratum shows the largest difference between sample and population shares (3.1 percentage points), followed by the economy issue area (2.4 percentage points). While there is some over-sampling of multi-issue IOs, these differences are small and should not distort our findings. Indeed, to rule out this possibility, we calculated sampling weights based on region and issue area using the IO distribution in the population and sample. The thus weighted data yield the same mean estimates as the unweighted data (see Table SI.9). This provides evidence that our sample is representative of the geographic and thematic distribution of IOs in the *relevant* population. Moreover, applying sampling weights, our calculated average IO authority, for example, can be interpreted as estimate of a population parameter.

As noted in the main text, we adjust the selection of IOs within some strata, especially where the sampling targets left us with one IO to choose from the population. In those cases, we deliberately sample better-known IOs to which scholars can connect and for which relevant questions can be answered, such as the evolution of authority of key international institutions (e.g IMF), instead of allowing pure chance to select the IOs. That some strata overrepresent better-known IOs does not mean that these IOs also have more authority. It is a matter of further empirical analysis whether IOs considered to be very salient do indeed convey high levels of authority. By contrast, it is well-known that regions and issue areas affect the distribution of IOs and their institutional design. We build on these insights by ensuring that the sampled IOs align to the regional and issue area distribution of the IO population list.

Still, we assess the implications of the over-representation of prominent IOs. To do so, we construct post-stratification weights based on IO prominence. To determine IO prominence, we collect the number of Google Scholar hits that an IO achieves in the period 2002-2011, using 500 hits as the threshold value of prominence (i.e. the 90th percentile of the Google Scholar hits distribution). We combined the prominence weights with the sampling weights from above and label the *combined weights*.

Table SI.9 shows the results for the means of overall and function-level authority. The first

Table SI.9: Weighted means for the IAD sample of 34 IOs and seven policy functions in 2010. Abbreviations: AS = agenda setting; RM = rule making; MO = monitoring; NI = norm interpretation; EN = enforcement; EV = evaluation; KG = knowledge generation; Authority = overall authority score. Standard deviation in square brackets. Standard error of mean in parentheses.

	Unweighted	Sampling weights	Combined weights
AS	0.50	0.50	0.42
	[0.41]	[0.41]	[0.44]
	(0.07)	(0.07)	(0.08)
RM	0.34	0.34	0.23
	[0.33]	[0.33]	[0.29]
	(0.06)	(0.06)	(0.05)
MO	0.17	0.17	0.13
	[0.22]	[0.23]	[0.18]
	(0.04)	(0.04)	(0.03)
NI	0.42	0.42	0.39
	[0.36]	[0.36]	[0.39]
	(0.06)	(0.06)	(0.07)
EN	0.17	0.16	0.08
	[0.27]	[0.26]	[0.18]
	(0.05)	(0.04)	(0.03)
EV	0.03	0.03	0.01
	[0.12]	[0.12]	[0.07]
	(0.02)	(0.02)	(0.01)
KG	0.28	0.27	0.30
	[0.44]	[0.44]	[0.43]
	(0.08)	(0.07)	(0.07)
N	238	238	238
Authority	0.27	0.27	0.22
	[0.18]	[0.17]	[0.17]
	(0.01)	(0.01)	(0.01)
N	34	34	34

column shows the seven policy functions, with the last row displaying the overall authority score. The second column reports the unweighted mean estimates. The third column shows the sampling weights that adjust for possible region and issue area imbalances. The fourth column displays mean estimates computed with the prominence and sampling weights (combined weights). Adjusting for region, issue area, and prominence, IOs display an average authority score of 0.22—that is 0.05 less than the unweighted average of IO authority (0.27). Applying weights also does not affect the general patterns of IO authority across functions. The two exceptions are knowledge generation (KG, being the third most authoritative function with weights). Moreover, with weights applied, enforcement ranks lower than monitoring.

C Variables in Dataset

Version 1.1 of the IAD data comprises 13 variables for a total of 1,694 observations (IO-years). The raw data, along with our normalized measure of IO authority, can be used to individually or jointly. Users of the data can tailor the data to their specific analytic needs. In the following, we briefly explain each variable that we code for each IO and year.

Variables contained in the dataset `'iadv11.dta'`

ioname Contains the official name of the IO, for example 'Food and Agriculture Organization'.

ioac IO acronym as used in research and IOs' public relations, for example 'FAO'.

iocode Numeric code of IO based on the IGO list of Pevehouse et al. (2004). The FAO's code, for example, is 1840. We adapt the use of the IGO codes in three cases. First, we assign the code '1830' to the European Community, the predecessor of the EU. Second, we assign the WTO's code of '4580' to the General Agreement on Tariffs and Trade. Third, we create a new code—'4115'—and assign it to the Shanghai Cooperation Organization (SCO). The SCO receives a new code as the organization was created after the publication of the IGO list in 2004/2005.

year Calendar year. The variable is first recorded for an IO in the year in which its constitutive treaty entered into force and the organization became operational. For example, the Treaty of Rome creating the European Communities (EC) entered into force in 1957 and thus the EC first appears in our data in 1957.

as Agenda Setting. The normalized score of an IO's agenda-setting authority, ranging between 0 (no authority) and 1 (full authority).

rm Rule Making. The normalized score of an IO's rule-making authority, ranging between 0 (no authority) and 1 (full authority).

mo Monitoring. The normalized score of an IO's monitoring authority, ranging between 0 (no authority) and 1 (full authority).

ni Norm Interpretation. The normalized score of an IO's norm-interpretation authority, ranging between 0 (no authority) and 1 (full authority).

en Enforcement. The normalized score of an IO's enforcement authority, ranging between 0 (no authority) and 1 (full authority).

ev Evaluation. The normalized score of an IO's evaluation authority, ranging between 0 (no authority) and 1 (full authority).

kg Knowledge Generation. The normalized score of an IO's knowledge-generating authority, ranging between 0 (no authority) and 1 (full authority).

authority This variable is the sum of the function-level authority scores per IO and year. Its (theoretical) value range is from 0 to 7.

nauth This variable is the normalized authority score, ranging between 0 (no authority) and 1 (full authority).

version This variable contains the version number of the current dataset.

D Contrast to MIA

In the main text we show that a direct comparison between our IAD measure and the MIA data is difficult to perform. Conceptually and methodologically, both datasets rarely overlap. Sufficient agreement is only found with respect to MIA's Dispute Settlement indicator and the IAD's Norm Interpretation function.

In addition, both projects code enforcement provisions. MIA incorporates a measure of sanctions for three decision stages/areas: fines against member states in financial arrears, curtailment of membership rights, and whether aggrieved state parties can apply retaliatory sanctions when other states fail to comply with rulings of dispute settlement bodies (i.e. remedies for non-compliance). By contrast, IAD considers a broad range of different sanctions that the IO could impose when states violate their substantive obligations and policies. These range from calls for corrective action, over financial and economic sanctions, to military intervention. MIA's Policy Making and IAD's Rule Making and Agenda Setting functions overlap conceptually. The coding, measurement, and aggregation, however, differ too substantially across MIA and IAD to allow for a direct contrast.

That both projects consider different sets of policy functions or decision areas can be clearly seen in Table SI.10 that contrasts the two projects at the indicator level. The entries in the first column show the MIA indicators, while the first row presents the IAD indicators (policy functions). Table SI.10 shows that the MIA indicators barely overlap with the IAD policy functions (indicated by empty cells).[4]

The exception is MIA's Dispute Settlement and the IAD's Norm Interpretation functions that are sufficiently comparable with regard to concept and measurement, because both projects use similar coding items and because MIA measures only the extent of delegation in dispute settlement, and not of pooling. In fact, both functions have similar means (NI: .345; sd = .341; DSM: .309, sd = .304), they cover the entire value range (min 0 to max 1), and correlate at .70.

Figure SI.3 compares both indicators in more detail for each IO included in both samples

[4]Abbreviations of column headings in Table SI.10 are: AS = Agenda Setting; RM = Rule Making; MO = Monitoring; NI = Norm Interpretation; EN = Enforcement; KG = Knowledge Generation; EV = Evaluation. Cell entries indicate the type of overlap. Empty cells indicate no overlap. Sufficient conceptual and measurement overlap is indicated by $+$. Conceptual overlap, but lacking measurement overlap is indicated by $*$.

Table SI.10: Differences and Similarities between IAD and MIA.

MIA/IAD	AS	RM	MO	NI	EN	KG	EV
Policy Making	$*$	$*$					
Dispute Settlement				$+$			
Assembly							
Executive							
Gen. Secretariat							
Consult. Bodies							
Members. Accession							
Members. Suspension							
Const. Reform							
Finance							

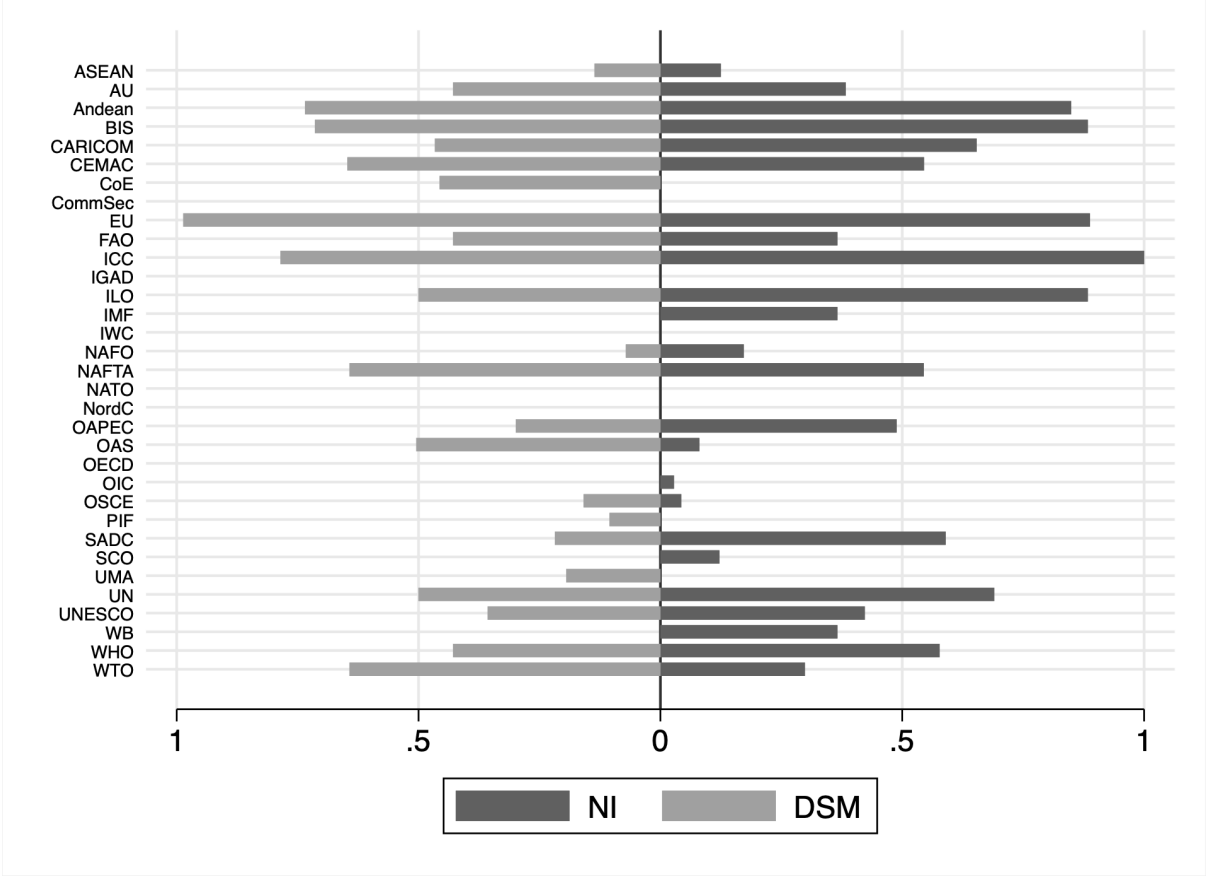

Figure SI.3: Norm interpretation (NI) from IAD and Dispute Settlement (DSM) from MIA compared. Comparison for the year 2010.

(N=33) in the year 2009. There is substantial agreement between the two datasets. In some cases, though, scores are different either because the IAD's norm interpretation function is broader, capturing more features of this policy function, or because both datasets use different measurements and apply different aggregation rules. For example, MIA assigns a positive value to the Council of Europe (CoE), while the norm interpretation function score zero. This is because MIA considers the European Court of Human Rights (ECtHR) to reflect the CoE's dispute settlement authority, whereas according to our IO selection criteria, the ECtHR is not included in the IAD sample as an independent organization (it is an emanation).

Possible uses

As argued and shown, MIA and the IAD capture different aspects of IO authority, with the former focusing mostly on internal governance arrangements of IOs and the latter on the external exercise of IO authority. In this section, we present descriptive statistics on MIA's pooling and delegation and the IAD's authority score. These can help scholars judge the respective features of each dataset and possible ways to combine certain indicators or subsets of each dataset.

Table SI.11 shows descriptive statistics for the IAD's single authority indicator and MIA's pooling and delegation scores.

Table SI.11: Means of Authority, Delegation, and Pooling

	Mean	SD	Min	Max	*N*
Authority	0.24	0.18	0.00	0.71	1411
Delegation	0.21	0.14	0.00	0.65	1411
Pooling	0.31	0.19	0.00	0.73	1411

The statistics have been calculated based on a sample of overlapping IOs. The unit of observation is the IO–year combination, totaling 1411 observations and covering the period from 1950 to 2010. Pooling has the largest mean (and the largest standard deviation), followed by the IAD's authority indicator. Delegation has a mean of .21, scoring lowest among the three different variables in Table SI.11.

In terms of changes over time, Hooghe et al. (2019, p. 38) inspect the growth patterns of delegation and pooling, respectively. Between 1975 (.18) and 2010 (.26), delegation grew in total by about 44.4%. Pooling grew too, but considerably weaker (14.3%). The average pooling

score was .28 in 1975 and .32 in 2010. The IAD's authority indicator started at .21 in 1975 and reached a score of .27 in 2010 (total percentage change of 28.6%). Averaging the growth of pooling and delegation yields a 30% increase. This growth pattern is slightly larger than the growth in the IAD's authority score. Yet, both point to a moderately increasing extent of IO authority over the period 1975–2010.

Table SI.12 shows the correlation between the IAD's authority and MIA's pooling and delegation.

Table SI.12: Correlation between Authority, Delegation, and Pooling

Variables	Authority	Delegation	Pooling
Authority	1.00		
Delegation score	0.59	1.00	
Pooling score	0.48	0.27	1.00

Authority correlates more strongly with delegation than with pooling. Yet, even with delegation, the correlation is rather moderate, indicating that the two measures, while in general terms moving into the same direction, capture different aspects of IO authority, such as the external exercise of IO authority (IAD) or the internal/constitutional degree of delegated authority to IOs (MIA).

Figure SI.4 presents scatter plots of authority with delegation and pooling, respectively. Authority and delegation correlate positively and more strongly with each other than authority and pooling. Points above the red dotted line denote IO–years with larger authority than delegation or pooling values. IOs below the red dotted line indicate IOs with higher delegation or pooling scores and lower authority scores.

Of note, several IO–years cluster at zero on the y-axis in both panels of Figure SI.4. While some IO–years with zero delegation achieve a positive authority score, the opposite does not hold. No IO with a zero pooling score has a positive authority score (bottom panel in Figure SI.4).

An IAD authority score of zero demarcates the baseline or reference category of our measure. It is defined as the non-institutionalization of a policy functions in legal texts (no authority by definition) or by a zero score on either autonomy or bindingness. IOs can deviate from that

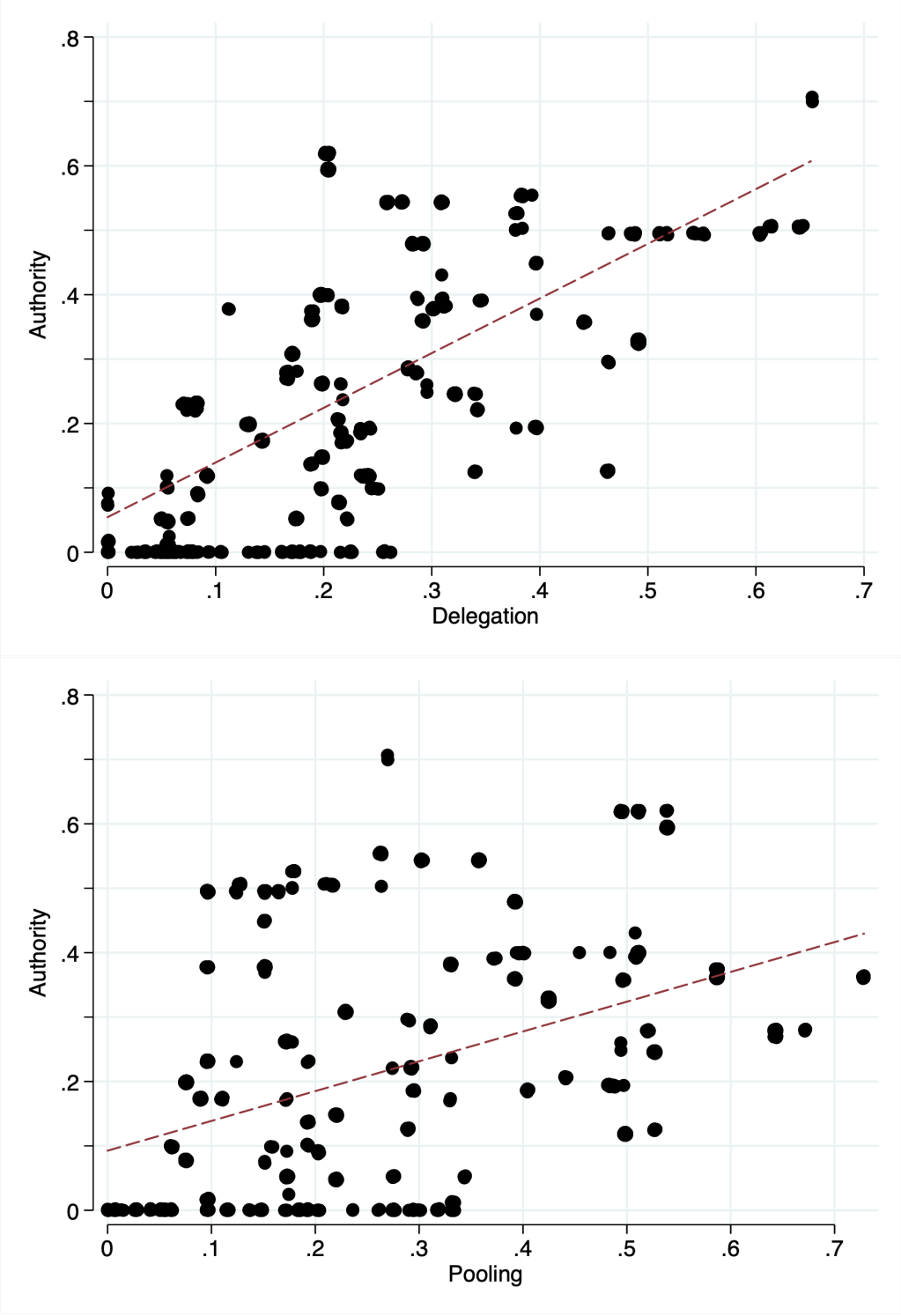

Figure SI.4: Scatterplots of Authority with Delegation (upper panel) and Pooling (lower panel). The dashed line indicates the line of best fit between the data. The data used for the scatter plot is the sample of overlapping IOs from 1950 to 2010 ($N = 1411$)

baseline only if a policy function is institutionalized and if its equipped with non-zero autonomy and bindingness values.

Figure SI.5 plots the IAD's authority score per IO along with MIA's delegation and pooling measures, respectively. The plots have been constructed based on the overlapping sample of IOs in the year 2010 ($N = 33$). We refrain from a direct comparison of IAD and MIA because the underlying concepts and methods to measure authority differ too strongly between the two datasets.

However, scholars wishing to combine the data or even just evaluate the distribution of external authority (IAD) with respect to internal/constitutional authority patterns (MIA) might find the data in Figure SI.5 useful for that purpose. For instance, the ILO appears to be an authoritative IO when judged by MIA's delegation and pooling scores with scores of 0.49 and 0.42, respectively. The ILO's internal and constitutional procedures delegate many competences to the secretariat and, with respect to pooling, majority decision making prevails. The ILO reaches an authority score of 0.32 based on the IAD. This indicates that the exercise of external authority over the seven policy functions is meaningful, but also that it lacks comparable levels of regulatory depth when contrasted to MIA's measures. Such contrasts between the allocation of authority across internal/constitutional processes and across processes of external, substantive policy making, for example, can stimulate new research questions about IOs' institutional design. Moreover, they can be used to generate new indicators for specific research questions, combining information from the IAD and MIA to more fully represent the distribution and exercise of authority across and within IOs over time.

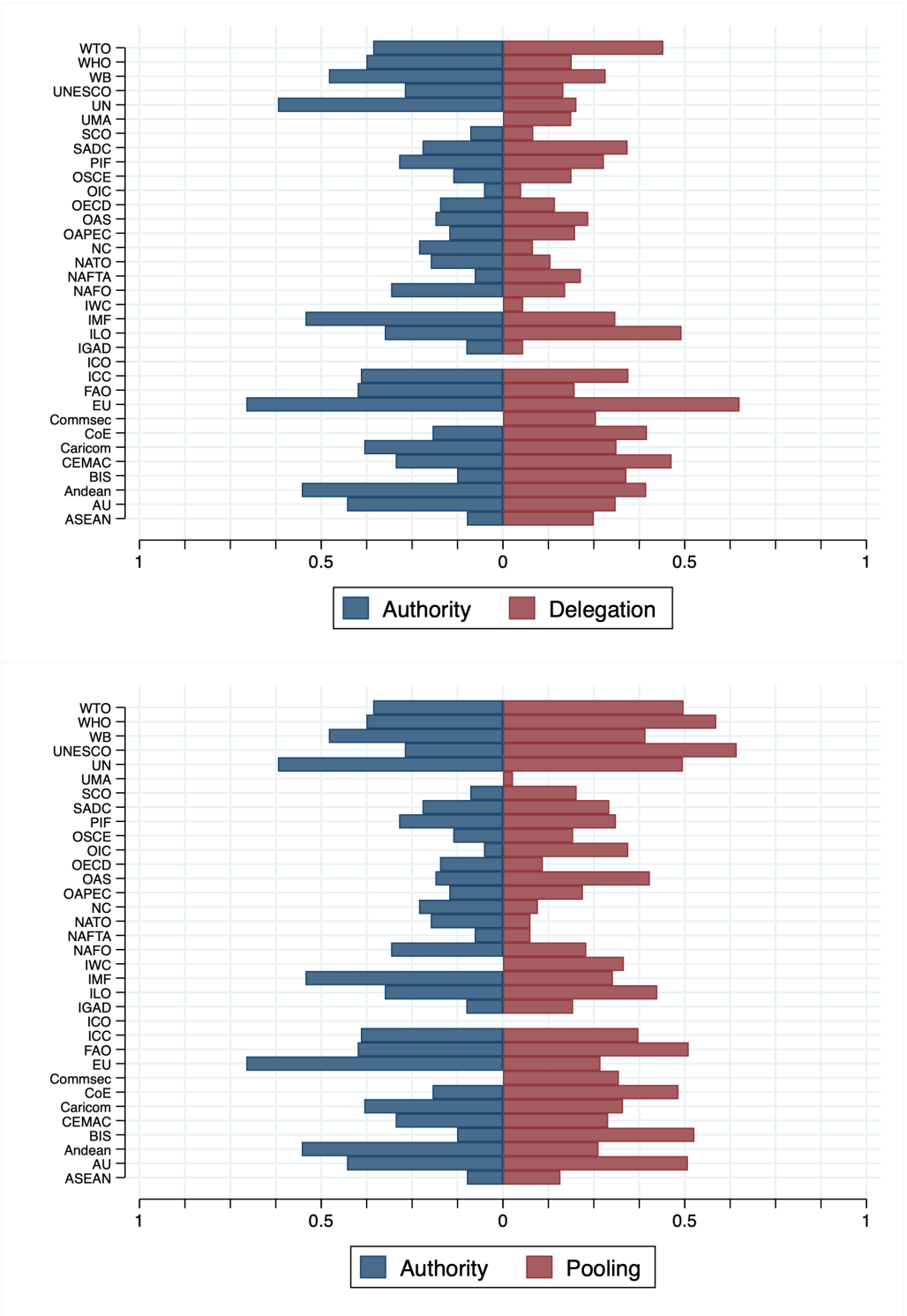

Figure SI.5: Comparing Authority to Delegation (upper panel) and Pooling scores (lower panel) for overlapping IOs in 2010.

References

Goertz G (2006) *Social Science Concept. A User's Guide*. Princeton, NJ: Princeton University Press.

Hooghe L, Lenz T and Marks G (2019) *A Theory of International Organization*. Oxford: Oxford University Press.

Pevehouse JC, Nordstrom T and Warnke K (2004) The COW-2 International Organizations Dataset Version 2.0. *Conflict Management and Peace Science* 21: 101–119.

Sommerer T and Tallberg J (2017) Transnational access to international organizations 1950-2010: A new data set. *International Studies Perspectives* 18(3): 247–266.