Mission completed?

Approaching international organizations from an incomplete contracts theory perspective using automated text analysis tools

Markus Gastinger (TU Dresden, Germany)
markus.gastinger@tu-dresden.de

Henning Schmidtke (GIGA German Institute for Global and Area Studies, Hamburg, Germany)
henning.schmidtke@giga-hamburg.de

ABSTRACT How "incomplete" are agreements setting up international organizations (IOs)? Does incompleteness vary across IOs and, if yes, how can this variation be explained? These are the questions at the center of this paper. We define incompleteness as language that is ambiguous and imprecise and, intentionally, leaves a wide margin of interpretation for the actors involved to construe its intended meaning after documents are first drawn up. Based on a review of political science literature, we show that this definition is part of most studies employing, implicitly or explicitly, incomplete contract theory. We then draw on the literature on the rational design of international institutions to transfer this perspective to the founding documents of IOs. We formulate hypotheses that understand incompleteness as an institutional design feature purposefully used by state actors and find empirical support for them across 72 IOs using a Beta regression model. Delegation and the number of member states increase incompleteness, which goes back to states' willingness to efficiently achieve specified objectives and the need for flexible language to reach compromise among independent actors. By contrast, pooling reduces incompleteness since it entails significant sovereignty costs. Beyond our contribution to the literature on the rational design of international institutions, our approach to measure incompleteness is generalizable to a wide variety of other contexts.

Paper prepared for the 12th Annual Conference of the Political Economy of International Organizations, Salzburg, February 7-9, 2019

(draft version – please do not cite without authors' permission)

"The only mention of [conditionality in lending] in the [Articles of Agreement] is a directive that the [International Monetary] Fund lend with '*adequate* safeguards,' but the term is never defined." (Barnett and Finnemore 2004: 57, italics added)

I. Introduction[1]

What a difference a word can make. In their seminal contribution on how international organizations (IOs) wield power through impersonal rules and the re-interpretation of mandates, Barnett and Finnemore (2004) show how the International Monetary Fund (IMF) exploited a vague mandate to gradually expand its reach into the domestic policies of member states running into balance-of-payments difficulties and turning to the Fund for financial assistance. In the Articles of Agreement, little information is included on what exact form its lending operations should take. There are various ways to make this point. One is to undertake a comprehensive reading of the entire document to come to a contextual understanding of what states intended, which would be a vain endeavor as there are no more specific provisions on IMF lending. Even if more ink would have been spilled, it is unclear whether this would have narrowed the scope of interpretation or provided additional inroads for IMF action. Take the quote at the beginning of this paper. We argue that the word "adequate" is not a random coincidence. Either states *genuinely wanted* to give the IMF wide latitude in this area, or the term goes back to purposeful *political compromises* to overcome conflicting positions translated into vague language – and the choice of words signals it.

International agreements are, at the end of the day, contracts. This allows casting light on them from a contractual perspective that is central not only to political science and

[1] For comments on earlier drafts of this paper, we are thankful to Anja Jetschke, Christian Kreuder-Sonnen, and the participants at the conference of the IR section of the German Political Science Association in Bremen in October 2017.

International Relations but also law and economics. Given that contracting underlies all kinds of economic interactions between rational actors, incomplete contract theory left a particularly lasting imprint on economics. As the culmination of this process, the Nobel Prize of 2016 was awarded to Oliver Hart and Bengt Holmström, two economists dedicating much of their professional careers to incomplete contracting. Their collaborative efforts are also our starting point to come to terms with this concept. We borrow the term "incompleteness" from Hart and Holmström (1987: 131–134) and define it as language that is "vague and ambiguous" and, intentionally, leaves a wide margin of interpretation for actors involved to construe its intended meaning in the aftermath of contracts being first drawn up. While the notion behind incompleteness, defined as equivocal language, is central to the legal profession (for instance, Endicott 2000; Mellinkoff 1963: 21–22) and may be as old as language itself, the formalization of its origins and consequences by economic theorists led to its spread across the social sciences.

The concept of incomplete contracting has grown vastly more prominent since the 1980s and early 1990s. To illustrate this point, we have searched for variations of the term "incomplete contracts" across three major scientific indexing services: Web of Science (WoS), Dimensions (D), and Google Scholar (GS).[2] The search query returns 1,218 results on WoS, 13,053 on D, and 18,000 on GS. However, all three platforms show that the term has entered the scientific discourse late but then quickly developed a widespread impact. According to WoS, the concept was invoked only 17 times until 1992 – but 1,201 times since then. Unsurprisingly, the lion's share accrues to economics (748), management (147), and business-related fields of study (203). But the concept has also become a household name in law (183), political science (42), and International Relations (25). As far as International

[2] In WoS and D, we used the query: "incomplete contract*". In GS, we searched for: "incomplete contract" OR "incomplete contracts" OR "incomplete contracting". Date of extraction: 1 January 2019. We also acknowledge that the general trend partly reflects the increasing number of indexed sources. However, we stand by our bigger point that incomplete contracting has risen in prominence and, today, is among the most widely used concepts in social science research.

Relations is concerned, this greatly underestimates the importance of the concept since the notion behind incomplete contracts has been mostly dealt with under the term "precision" (see below).[3] In D, the term incomplete contracting pops up 326 times up until 1992 and 12,727 times since then. The distribution across disciplines is broadly comparable to the WoS, with applied economics (4,151) and "business and management" (1,919) topping the list. However, the numbers for law (867), "policy and administration" (795), and political science (543) are testament to the concept's reach beyond economics.[4] While GS does not allow to distinguish various fields of study, it impressively corroborates the overall trend. Until 1992, incomplete contracting has been mentioned 1,110 times. It has been used almost 17,000 times since then. However, better strategies to empirically measure incompleteness are sorely needed.

We use automated text analysis tools to measure incompleteness. This allows mass processing of large swaths of data at greatly reduced costs and enhances reliability and replicability by removing subjective human judgment as a source of error. Drawing on political science and International Relations literature, we identify a number of words that are inherently incomplete, such as "adequate", "necessary", or "essential". These terms can be used in a wide variety of contexts but always indicate that the drafters of any given document wanted to leave the door open for subsequent (political) interpretation. While the use of these terms could also go back to sloppy drafting or plain "error", our argument rests on the assumption that they are purposeful political decisions manifesting themselves in vague language. If this were not the case, variation in the texts we study should follow a random pattern and statistically significant correlations with other variables of interest are unlikely to occur. Given that our empirical analysis of international organizations' (IOs) founding treaties

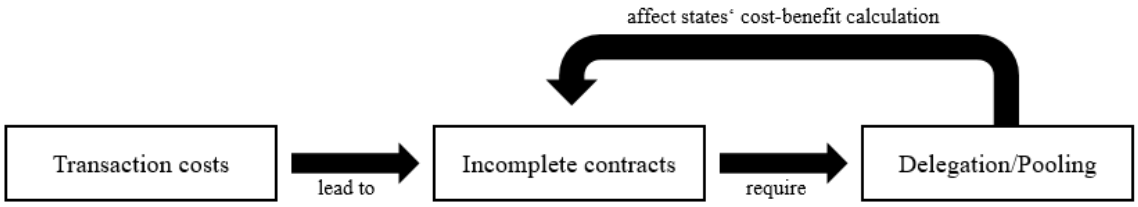
[3] We do prefer to stick with the term "incomplete contracts" in this paper, even if most aligned with the field of International Relations, in order not to lose the link to other disciplines and to underline the wider applicability of our approach.
[4] D has no separate listing for International Relations.

does yield statistically significant results, we have some confidence in our approach validly capturing incompleteness.

Specifically, we find that the incompleteness of IO treaties is explained by the level of authority delegated to IOs, which indicates a willingness among states to efficiently achieve the specified objectives and reluctance to tie their agents' hands by resorting to inflexible language. Moreover, the number of IO member states increases the need for flexible language to reach a compromise among all the actors involved in the negotiations and results in incomplete treaties. At the same time, the pooling of political authority in collective decision-making bodies reduces incompleteness since it entails significant sovereignty costs for states, thereby providing them with a powerful incentive to minimize ambiguity.

In past research, incompleteness has primarily been used to explain delegation and pooling, which can be viewed as "solutions to the problem of 'incomplete contracting'" (Moravcsik 1998: 73). Incomplete contracts were often simply assumed to exist and considered impossible to avoid. We accept this premise only to a degree and prefer to treat incompleteness as "just another design option" available to states. Put simply, states need to come to a decision whether to accept incomplete contracts at the same time that they balance other design options such as the extent of delegating authority or pooling sovereignty. Given these other design options, their cost-benefit calculation of incompleteness will change and resulting contracts will be more or less incomplete (see Figure 1). Above and beyond our contribution to the literature on the rational design of international institutions (Koremenos *et al.* 2001), our paper contributes to a growing number of studies exploring the potential of computational applications using text as data. While some readers may object to our approach on the grounds that it is too simple to merit serious attention, Grimmer and Stewart (2013: 270) actually suggest that "models that are less sophisticated in the use of language may provide more useful analysis of texts".

Figure 2: Schematic view of this paper‘s argument



Notes: Actors defined by bounded rationality incur transactions costs in the process of drafting contracts (thinking costs; decision-making/bargaining costs; drafting/writing costs; enforcement costs). To overcome incomplete contracting, states opt to delegate authority or pool sovereignty in IOs to facilitate subsequent cooperation (i.e. more readily establish authoritative interpretations of intentionally ambiguous provisions, which is the focus of this paper, and filling gaps that may or may have not been left out in the contract intentionally). The prevalent argument in the literature is that all contracts are incomplete. This, however, would logically imply that all IOs witness the same, or at least similar, levels of delegation and pooling, which is not in tune with the reality of IO constitutions. Consequently, rather than using incomplete contracts to explain delegation and pooling, we submit that states’ decision to accept incompleteness is partly determined by their concurrent decision to delegate authority or pool sovereignty.

The remainder of this paper is structured as follows. Section II presents our theoretical argument. This entails two points. First, we review incompleteness in the political science and International Relations literature and survey past attempts to measure it empirically. Second, we will deduce hypotheses on what explains incompleteness in the case of IO constitutions. Section III introduces our sample and operationalization of variables. Section IV contains the empirical results. Section V concludes.

II. Theoretical argument

We begin this section with a literature review of contributions on incomplete contracts with a focus on how they are measured empirically. This further clarifies our own understanding of the concept. While incompleteness – and our approach for measuring it – is transferable to a wide array of different areas, the factors explaining it will require theorizing tailored to the specific context. While we do not rule out that a grand theory of incompleteness may be attainable in the future, we are interested here specifically in the question of what explains incompleteness in IO constitutions, that is, the very treaties negotiated among states establishing IOs as common institutional frameworks. We hypothesize that delegation, pooling, and the number of member states determine incompleteness in these cases.

Incomplete contracts, political science, and the difficulties of empirical measurement

In political science, incompleteness is imbued with both negative and positive effects. Negatively, it complicates the assessment of whether states comply with their international commitments (Abbott *et al.* 2000: 414; Chayes and Chayes 1993: 188–192) and enables self-interested interpretations by states or IOs (Barnett and Finnemore 2004; Linos and Pegram 2016). Putting incompleteness in a more positive light, it increases flexibility, reduces transaction costs by allowing negotiations to conclude absent "complete" agreements, and enables endogenous institutional change (Best 2012; Cooley and Spruyt 2009; Farrell and Héritier 2007; Héritier 2012; Koremenos 2013: 143). Whether positive or negative, given the bounded rationality of actors, their inability to foresee each and every future contingency and costs associated with drafting complete agreements, all contracts are "*unavoidably incomplete*" (Williamson 2000: 601; emphasis original). Still, some contracts are more

incomplete than others (Mattli and Stone Sweet 2012: 8) with the ability to discern differences only hampered by the challenges connected to its empirical measurement.

Incompleteness is indeed a highly elusive phenomenon to measure empirically. In this section, we survey the extant political-science literature for the notion of incomplete contracts with a focus on how the challenge of empirically measuring the phenomenon has been approached. Importantly, we are not only looking for contributions in which the term "incomplete contract" has been used explicitly. Rather, we aim to present a more comprehensive review by including contributions that, while not using the term itself, take aim at essentially the same phenomenon, albeit from various perspectives and against various backdrops. Table 1 summarizes the key attempts at measuring incompleteness, distinguishing manual from automated approaches, and categorizing them in terms of validity, that is, how accurately they promise to measure incompleteness.

Table 1: Approaches to measuring incompleteness.

	Manual	Automated
High validity	Rules vs. standards (Ehrlich and Posner 1974; Abbott *et al.* 2000; Koremenos 2016)	-
Medium validity	Shall vs. may (Linos and Pegram 2016); textual ambiguity (Best 2012)	Dictionary of incomplete words (Gastinger and Schmidtke); standard clauses (Manger and Peinhardt 2017); scope and depth (Dür and Lechner 2019)
Low validity	-	Word count (Huber *et al.* 2001)

Source: Authors' own compilation.

While not mentioning incomplete contracting explicitly, it can be readily linked to the notion of *precision* in international law introduced by Abbott and colleagues (2000: 412–415). A precise rule "specifies clearly and unambiguously what is expected of a state or other actor

[and] narrows the scope for reasonable interpretation“ (Abbott *et al.* 2000: 412). It thereby increases predictability in the area to which it applies by avoiding self-interested interpretation. Where interpretative powers have been delegated to a court or arbitration body, precedent will establish increasingly precise rules even where treaties are initially vague (Schmidt 2018). However, Abbott and colleagues recognize that operationalizing precision is "difficult" (Abbott *et al.* 2000: fn 35). Following Ehrlich and Posner (1974: 261), who themselves acknowledge "the absence of a readily identifiable empirical counterpart to the concept of precision", Abbott and colleagues offer five broad indicators, ranging from specific "rules" to general "standards", against which incompleteness can be judged qualitatively. While this leads to measurement with high accuracy if assessed by subject-matter experts, this approach is difficult to replicate, highly labor-intensive, and (at least currently) impossible to implement using automated methods. Koremenos (2016: 161) acknowledges that coding precision often requires a "judgement call" and that, despite the prominence of incompleteness in political science, law, and economics, "no quantifiable guidelines have been elaborated [yet]".[5]

A related approach is showcased by Linos and Pegram (2016), who determine "flexibility" in agreement language by relying on vagueness. To assess vagueness, the authors initially seem to intuitively use an approach similar to the one we advocate in this paper by looking for words such as "*substantial* quantities" (2016: 593, italics added). Moreover, they include options and caveats in their list of linguistic techniques indicating flexibility, which we treat no differently from other parts of agreement texts as they make it more or less incomplete depending on the specific wording. More importantly, Linos and Pegram (2016: 602) concede "some subjectivity" in their classification and seem to rely, mainly, on the difference between the modal verbs *shall* and *may* to distinguish strong from flexible

[5] Koremenos relies on manual coding of well-trained student assistants, with the assistance of international law experts in case of disagreements. We provide further details on her approach in the research design section.

provisions. While this is a legitimate heuristic device to cognitively reduce complexity and coherently fit text into pre-defined categories, it risks conflating "precision" with "obligation" as defined by Abbott and colleagues (2000). These are two distinct concepts that should be kept separate.[6] We, therefore, deem their approach somewhat less valid than the other manual procedures as far as assessing incompleteness is concerned. Their central finding that vague language is no different from omitting issues altogether, however, underlines that more scholarly effort on how to identify and measure vagueness has far-reaching consequences.

Jacqueline Best also begins with a definition of ambiguity close to the approach that we develop more fully in this paper. She explains: "Textual ambiguities, such as the use of an open-ended term like 'necessary' or 'critical,' create considerable room for an expansive interpretation, as a growing range of issues get included in that elastic term" (Best 2012: 680). However, she uses textual ambiguity very sparsely in the empirical section of her paper, drawing instead heavily on her second face of ambiguity – institutional ambiguity – which finds expression in the vagueness of daily institutional routine practiced by international bureaucrats rather than written words. Textual ambiguity alone cannot reach the same level of validity as approaches relying on additional contextual factors to assess incompleteness, which is why we give it medium validity. This also seems fitting given that we rely on essentially the same approach in this paper and claim medium validity compared to more contextual in-depth approaches.

[6] One could use "shall" and "may" (including best-endeavors formulations) to develop an automated approach similar to the one we develop for incomplete words. However, we do not deem that a very promising route. A formulation such as "the parties shall *substantially* increase exports" is less precise than "the parties *endeavor* to increase exports by 20 percent". The former "obligation" may arguably already be satisfied by a five percent increase (how much more is substantially more?), making it difficult for any party to seek redress even if dispute settlement provisions are included. Under the latter formulation, the same increase would cast serious doubt on the parties acting in good faith under the terms specified in the agreement and entail at least reputational costs (cf. the phenomenon of "vagueness contagion", which we introduce in the research design section below).

There is a second dimension to incompleteness that we have not yet covered. Already Hart and Holmström (1987: 131–134) noted that incompleteness can also be the result of missing provisions ("gaps") in a contract.[7] Jean Tirole, another Nobel prize winner, similarly acknowledges these two dimensions when writing that "[incomplete] contracts *are vague or* silent on a number of key features" (1999: 471; italics added), as do political scientists such as Epstein and O'Halloran (1999: 37). This second dimension can be gauged through the scope and depth of a contract (Hooghe and Marks 2009: fn 3; Lenz *et al.* 2014; Mattli and Stone Sweet 2012: 8). One may suggest that the mere length of a document gives us some indication about this sort of incompleteness. For example, Huber *et al.* (2001: 336–338) argue, from a principal-agent perspective, that detailed acts of delegation limit the discretion of agents. Whether an act is detailed is measured by the total number of words since more extended contracts will, on average, constrain agent actions more. While Huber *et al.* thereby suggest an automated approach to measure incompleteness, this is just about the crudest proxy that we can imagine. Moreover, longer texts can also make documents *more* incomplete if the language used is rife with ambiguity, which gives agents seeking task expansion a foothold in a given area. Consequently, we include their approach in the bottom right of Table 1.

The study by Manger and Peinhardt (2017) offers a more refined approach to measure the second dimension of incompleteness. Drawing on Abbott and colleagues (2000), they measure the "precision" of almost 3,000 bilateral investment treaties by constructing a corpus of standard clauses. The more of these clauses are included, the more "precise" Manger and Peinhardt consider the treaty to be. To our mind, what is measured here is the second dimension of incompleteness as proxied by the scope of bilateral investment treaties rather than precision. This definitional differences apart, they develop an interesting approach that should yield more valid results than a simple word count. The biggest downside to their

[7] Jean Tirole, another Nobel prize winner, makes the same argument when writing that "[incomplete] contracts *are vague or* silent on a number of key features" (1999: 471 italics added).

approach is that it can be applied *only* to the analysis of bilateral investment treaties. The same is not true for more sophisticated text analysis techniques, which can computationally gauge the scope and depth of agreements (Dür and Lechner 2019). While this is promising in terms of using automated text analysis tools to measure the second dimension of incompleteness, it does not render unnecessary a reasonably valid strategy to measure what we have defined as its first dimension. Moreover, the two dimensions are closely interlinked. The mere fact that a greater number of issues is covered by a contract does not necessarily make it less ambiguous. This will, after all, still heavily depend on the precision of the language used.

Explaining incompleteness in IO constitutions

While we aim to develop a universal approach to measuring incompleteness in this paper, the factors *explaining* incompleteness require tailoring to specific contexts. In this section, we take three pertinent variables in the literature on the rational design of international institutions – delegation and pooling of authority, and the number of member states – to explain incompleteness in IO constitutions. Moreover, since there is nothing close to a "true" value for incompleteness that we can compare our measurements to (hence this paper), we also rely on these hypotheses as a first testing ground. If our approach is misguided, it should yield random data and we should not get statistically significant results in the specified direction. If the hypotheses are confirmed, this would give us some confidence in measuring (the first dimension of) incompleteness.

Delegation can be defined as the "conditional grant of authority from a *principal* to an *agent* that empowers the latter to act on behalf of the former" (Hawkins *et al.* 2006: 7; emphases original). In International Relations, the member states of any given IO are its "collective" principal (Nielson and Tierney 2003) while the IO's secretariat and other

supranational bodies are their agents. Agents not only serve to overcome problems of incomplete contracts, making delegation "often the best way" to deal with incompleteness (Abbott and Snidal 2000: 433; see also Abbott and Snidal 2010: 324; Garrett 1992: 557). As Abbott and colleagues (2000: 415) point out, imprecise rules coupled with extensive delegation also enhance discretion for international bodies. If delegation underlines states' willingness to efficiently reach the objectives that the IO has been mandated to achieve, precise contracts risk counteracting these benefits by tying the agent's hands (Hawkins *et al.* 2006: 27–28). While incomplete contracting is neither necessary nor sufficient for delegation (Moravcsik 1993: 509), we formulate our first hypothesis as follows:

> *H_1 (delegation hypothesis): The more authority member states delegate to IO secretariats, the more incomplete treaties are.*

Pooling is the transfer of authority from member states to a collective state body. It increases when (binding) decisions can be taken by some form of (super-)majority rather than only unanimity (Blake and Payton 2015). By upholding the unanimity requirement, states tie *their own* hands and make the achievement of goals harder. As Lake (2007) has pointed out, the primary difference between delegation and pooling is the underlying strategic problem. The strategic problems connected to delegation are agency losses, that is, international secretariats using their delegated powers to pursue their own preferences (see also Tallberg 2002). This problem can be, at least partially, overcome by instituting more stringent mechanisms of control, ex-ante in the form of more thorough agent screening and selection or ex-post through increased monitoring or sanctioning mechanisms (Pollack 1997; Thatcher and Stone Sweet 2002). But there is no institutional solution for the strategic problem associated with pooling, which is collective decision-making. Since pooling means surrendering national veto, states being outvoted by a coalition of other states will be forced to accept a new status quo.

Therefore, "[m]uch of the concern with IOs is really about pooling rather than [delegation]" (Lake 2007: 220). Incomplete treaties will thus give rise to concerns about expanding sovereignty costs. This leads to our second hypothesis:

> *H_2 (pooling hypothesis): The more member states pool sovereignty in IOs, the less incomplete treaties are.*

Finally, IO treaties need to be negotiated among member states before being established. Since state preferences usually do not overlap perfectly, these negotiations become more difficult with an increasing number of states. States can use incompleteness as a strategic device to overcome diverging preferences, not by eliminating ambiguity but by accommodating it. Vague language enables states with heterogeneous preferences to interpret the treaty text in their favor, which facilitates the conclusion of negotiations. Historical case studies on the creation of IOs – including the International Monetary Fund, the World Bank, or the United Nations – highlight this effect (Barnett and Finnemore 2004; Best 2005; Helleiner 2014). Similarly, Koremenos (2016: 173) argues that imprecision increases with the number of member states since bargaining problems increase. Although she refutes the hypothesis based on the examination of her dataset of 234 international agreements, we subject it to further empirical analysis in our paper:

> *H_3 (bargaining hypothesis): The greater the number of member states negotiating the foundation of an IO, the more incomplete treaties are.*

III. Research design

This section explains the IO sample, outlines the approach used to gauge incompleteness in IO treaties, and operationalizes the dependent and independent variables.

Measuring incompleteness

Our analysis uses a sample of 72 IOs that pass a minimal threshold of formal organization originally assembled by Hooghe and colleagues. IOs in the sample have (1) an entry in the Correlates of War dataset, (2) a distinct physical location or website, (3) a formal structure such as an administrative body; (4) at least 50 permanent staff, (5) a written founding treaty or convention and (6) a decision body meeting at least every year (Hooghe *et al.* 2017: 16). The sample includes a broad variety of IOs from different world regions and issue areas. It includes the full spectrum of levels of membership, pooled and delegated authority, thereby maximizing the variance of our explanatory variables.[8]

For our analysis, we take the original founding treaties. Coming by these treaties was more demanding than we originally anticipated since they are not always easily available. To overcome this problem, we combined three different data sources. First, we used the United Nations Treaty Collection where some founding treaties are deposited (United Nations 2018). Second, we have obtained some documents from IO secretariats directly. Finally, we have asked Hooghe and colleagues for support and gratefully received the documents that we could not retrieve through any other means (for an overview see Appendix A1). The founding treaties were in various data formats and we have converted all of them into simple text

[8] Compared to the dataset by Hooghe and colleagues, our sample does not include the Benelux Community, the Commonwealth of Nations, the International Criminal Police Organization, and the second generation of the East African Community. These organizations either do not have formal founding treaties or we could not find official English-language versions.

format. Since the conversion process was sometimes erroneous and our approach relies on correct spelling, we have engaged in the laborious task of reading all converted files and manually typed up sections that proved impossible to convert from the original files. The resulting set of machine-readable IO founding treaties constitutes a unique data source for future research, which we shall make available to other researchers at the time this paper is published.[9]

In our conception, incompleteness goes back to one empirical referent – vague and imprecise language. Best (2012: 677, 680) mentions "necessary", "critical", and "essential" as ambiguous terms conveying multiple meanings. Linos and Pegram (2016: 593) illustrate their approach by specifically mentioning "substantial". Barnett and Finnemore (2004: 57) contribute the word "adequate".[10] These words can be grouped into two broader categories. First, words converging around "adequate" and "necessary". Second, those that converge around "critical", "essential", and "substantial". Starting from those two categories, we added several synonyms included in *The Oxford Thesaurus of English*. Ultimately, our dictionary of incompleteness consists of 12 ambiguous words.[11] By drawing on the International Relations literature to develop our dictionary, we increase the chances of capturing words indicating incompleteness in the case of IO constitutions. Nevertheless, we expect this dictionary to perform reasonably well across many areas since the identified words appear to us to universally indicate vagueness in most contexts.

In compiling the dictionary, we do not go as far as Koremenos (2016: 160–161) and refrain from including general terms such as "opinion", "environment", or "development". It

[9] We only use the main body of treaties for our analysis, excluding annexes, protocols, and other types of texts clearly set apart from the main body. The reason for this is twofold. Theoretically, annexes and protocols can include provisions that do not apply to all IO members and contain highly technical details, which are generally subject to relatively lower levels of hard ("linguistic") bargaining. Practically, it would have been impossible to correct the spelling of all these supplementary text sources, which run into thousands of additional pages.
[10] We should note that the quotation at the beginning of this paper was introduced into the Articles of Agreement only in a later amendment and is not included in our dataset. Nevertheless, this leaves the general argument developed in this paper unaffected.
[11] The full dictionary is available in the R-Script, which we shall make available at the point of publication.

is correct that language has inherently fuzzy boundaries, as philosophers of language have pointed out (Barker 2002; Keefe 2000). One could include words across all lexical categories in a dictionary of incompleteness, including gradable adjectives (for instance, big, small, high, low) and even nouns such as "heap", as the sorites paradox teaches us.[12] However, we prefer to stick closer to words that are more likely to capture the genuine intention to provide greater flexibility to IOs or purposeful political compromises, rather than the inherent ambiguity of language itself.

We use the statistical software *R* and draw on functions included in its base package to count the number incomplete sentences per IO treaty. We follow the insight of linguists that vagueness is "often contagious, in the sense that complex expressions built up from vague predicates are often themselves vague as a result" (Barker 2010: 1038). Consequently, we do not simply count the number of occurrences of incomplete terms in any given document.[13] Instead, we first break up documents into textual sub-units because this form of "vagueness contagion" operates on the level of sentences and not entire documents. To unitize our data, we follow the exogenous approach by Däubler *et al.* (2012) and use common punctuation marks. We include full stops, which mark the end of a sentence that is a complete statement; and semicolons and colons, which both usually indicate a break between main clauses and are used in lists.[14] After dropping sentences of four words or fewer to reduce noise in the data, we are left with around 17,000 sentences across 72 IO treaties. Whenever at least one of the words included in our dictionary is part of the sentence, we code the whole sentence as incomplete. As a first plausibility probe, we randomly selected 200 incomplete sentences and checked that the incomplete words in our dictionary, in fact, lead to considerable margins of

[12] The sorites paradox (or "paradox of the heap") is attributed to the Greek philosopher Eubulides of Miletus and, in short, begins with the observation that one grain of sand is not a heap. If that is true, then also two grains of sand cannot be a heap, nor three, nor four etc. In the end, no amount of sand ever passes as a heap. This paradox holds a lesson in logical fallacies just as much as vagueness, since it will always remain ambiguous how many grains of sand are necessary for a heap to form.

[13] However, we do adopt this approach as part of our robustness checks (see below and the Appendix).

[14] Exclamation and question marks can also end sentences. However, they do not occur in IO constitutions.

interpretation. To arrive at the overall level of incompleteness, we computed the share of incomplete sentences over all sentences for each IO. Our dependent variable thus is bound between zero and one.

Explanatory variables

We analyze the determinants of incompleteness by resorting to three variables. First, we hypothesize that the level of *pooling* of formal authority in intergovernmental IO-bodies decreases incompleteness. Pooling measures how much authority states have transferred to collective IO member state bodies and decided to move away from the unanimity requirement. It consists of three elements: the rules under which decisions are made (unanimity or some form of majority); whether decisions need to be ratified or enter into force without ratification by member states; and whether IO decisions are nonbinding, partially binding, or binding.

Second, we expect incompleteness to rise with the *delegation* of authority to IO bodies, that is, international bureaucracies, parliaments, and judicial bodies. Delegation measures how much formal authority states have granted to IOs. It is measured across nine indicators, including the IO secretariat's executive functions or involvement in policy initiation, budget drafting, and constitutional revision. Data for both variables across all IOs in our sample are provided by Hooghe *et al.* (2017) who have coded IO treaties and other formal rules and procedures on institutional design. We use these individual variables of pooled and delegated authority for the year of IO creation.

Finally, we examine the effect of *membership* on the basis of a simple count variable on the number of funding IO members. These data are mainly taken from the COW-International Organizations Dataset (Pevehouse *et al.* 2004). We have added missing values based on our own count of signatory states listed in founding treaties.

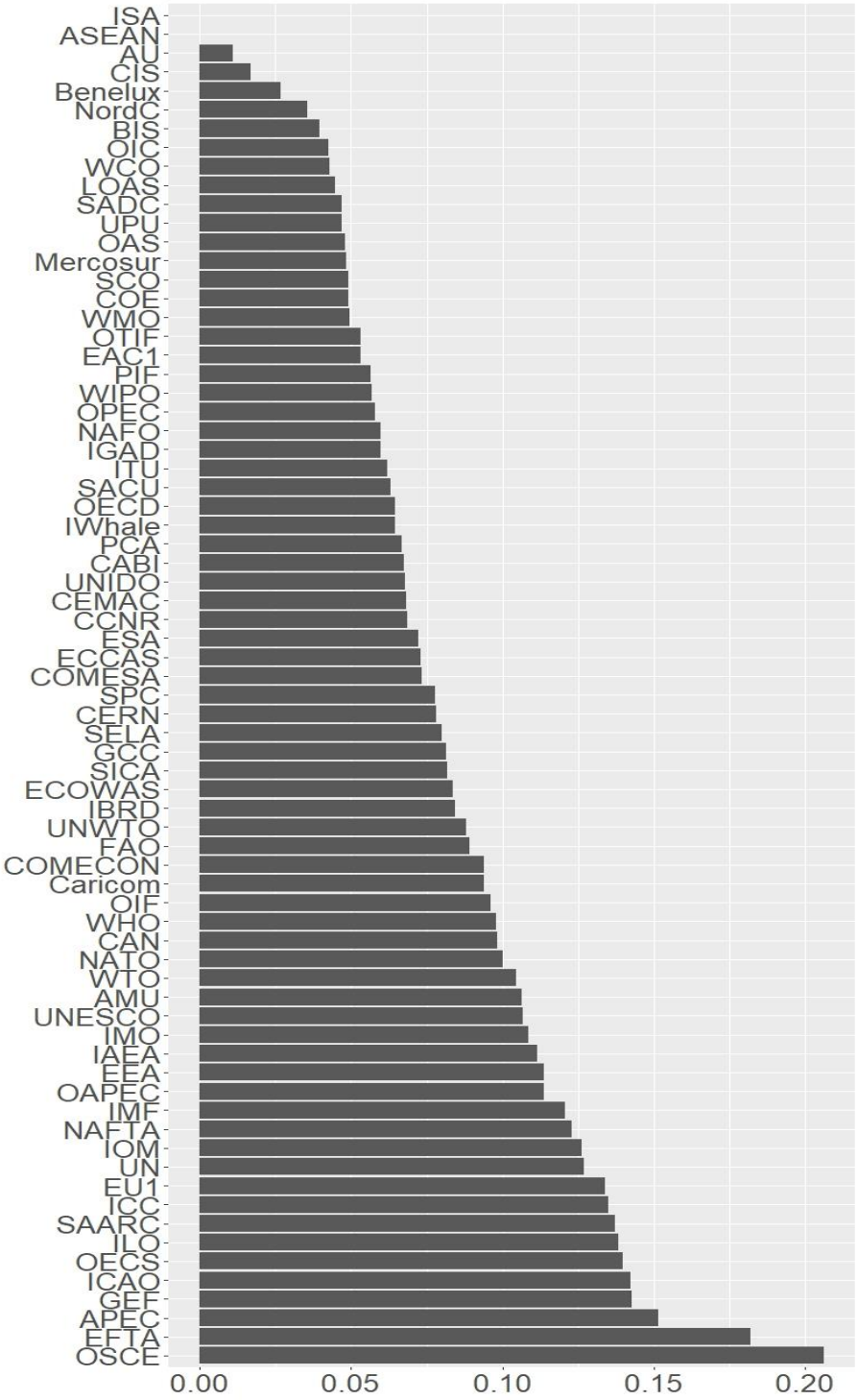
IV. Results

How incomplete are IO constitutions? Does incompleteness vary significantly across IOs? If yes, what explains this variation? Our empirical analysis of these questions proceeds in two steps. First, we present descriptive findings on IO treaty incompleteness and map central patterns. Second, we employ a Beta regression model to examine the explanatory power of our explanatory variables and to probe into the determinants of incompleteness.

Patterns of incompleteness

IOs display extensive variation of incompleteness across founding treaties. Figure 2 displays incompleteness using a ratio of incomplete sentences to the overall number of sentences in a given treaty.

Figure 2: Incompleteness across IOs



The figure reveals stark differences for the IOs in our sample. With more than 20 percent of sentences, the founding treaty of the Organization for Security and Co-operation in Europe (OSCE) stands out as a highly incomplete contract. On the other side of the distribution, we

find that the International Seabed Authority (ISA) and the Association of Southeast Asian Nations (ASEAN) operate on treaties that do not contain incomplete sentences. Interestingly, we find, with the OSCE and ASEAN, two organizations on both ends of the distribution, which have over time transformed from informal clubs into formalized IOs. This seems to suggest that incompleteness is not strongly driven by path-dependence or previous levels of incompleteness. The mean share of incomplete sentences is around eight percent and the median of the distribution is around seven percent. Overall, the data reveal substantial variation.

Previous research on the incompleteness of IO contracts has shown that the levels of incompleteness vary across issue areas (Goldstein *et al.* 2000; Koremenos 2016). Our data only partially confirm this assessment.

Figure 3: Incompleteness across issue areas

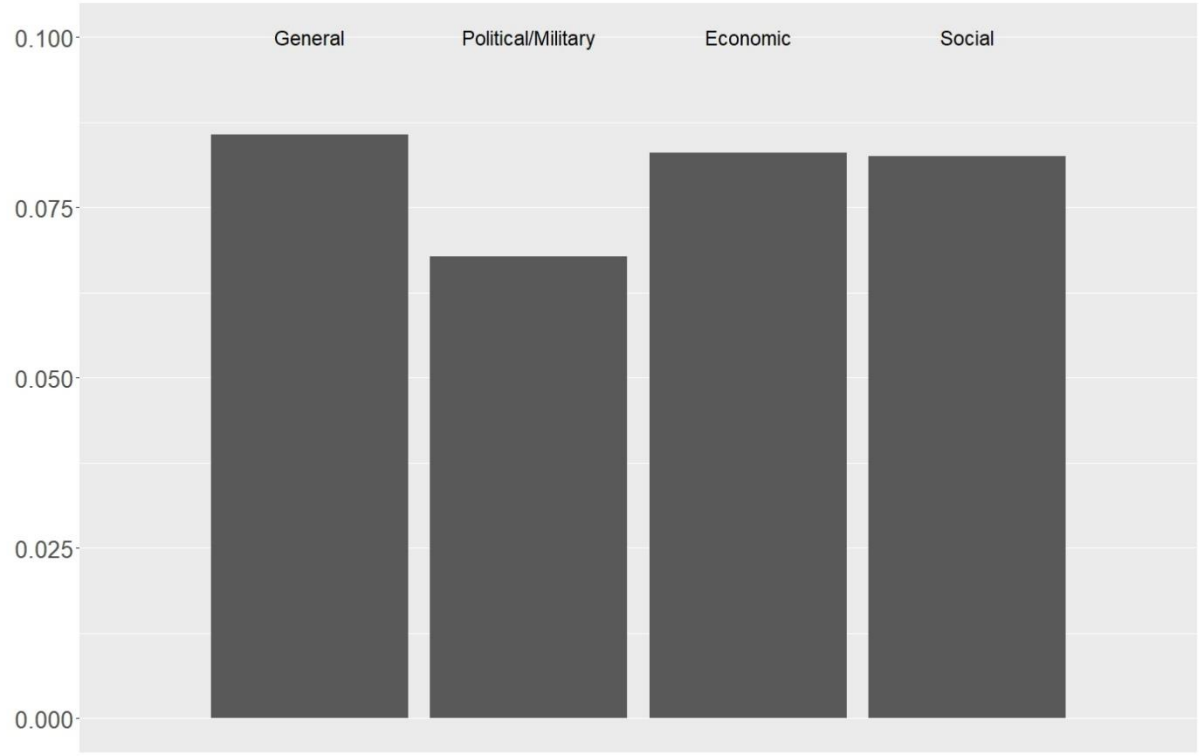
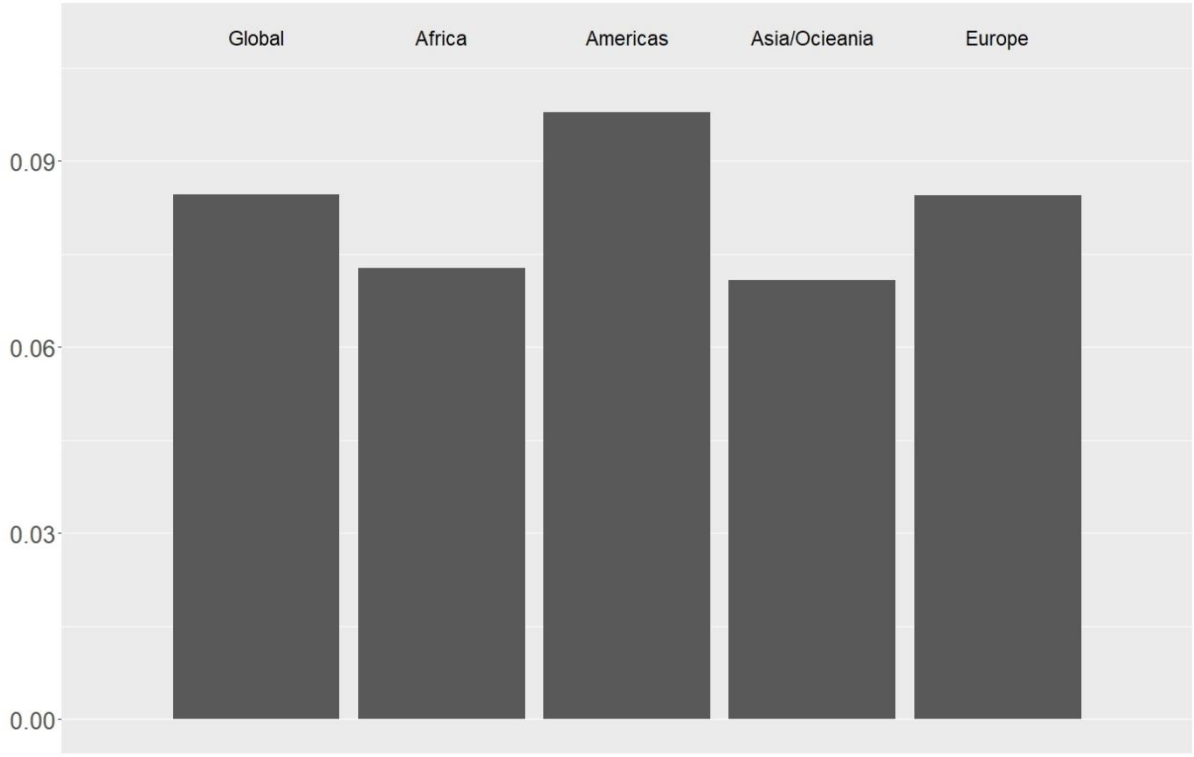


Figure 3 shows the incompleteness of founding treaties across four key issue areas of global governance (Volgy *et al.* 2008) where the scores represent the mean shares of incomplete sentences by issue area. Founding treaties of general multi-purpose organizations, including, for instance, the United Nations (UN), are the most incomplete (mean = nine percent). Yet, social (mean = eight percent) and economic-purpose organizations (mean = eight percent), such as the International Labour Organization (ILO) or European Union (EU), operate on similarly incomplete mandates. The lowest level of incompleteness is present for political/military IOs, such as the North Atlantic Treaty Organisation (NATO). This seems to imply that the structure of underlying cooperation problems founding states faced during the creation of IOs is a less important driving force of incompleteness than much of the rational design literature suggests (Koremenos 2016).

By contrast, incompleteness varies considerably across world regions (Figure 4). Incompleteness is most extensive in American IOs (mean = 10 percent) and less pronounced in Asian (mean = seven percent) and African IOs (mean = seven percent). Global (mean = eight percent) and European (mean = eight percent) IOs display a medium range of incompleteness.

Figure 4: Incompleteness across regions



Multivariate analysis

What explains this variation in the incompleteness of IO founding treaties? We test our hypotheses using Beta regression.[15] Our choice of a Beta regression is driven by both theoretical and methodological considerations. First, we focus on the *ratio* of incomplete sentences to all sentences in IO treaties because we are interested in the overall scope for interpretation a treaty offers to both member states and international bureaucrats. Second, given this type of data of our dependent variable, Beta regression is the most appropriate model for our analysis (Ferrari and Cribari-Neto 2004). This model is based on the assumption that the dependent variable is beta-distributed, assuming values in a standard unit

[15] We use the R software-package 'betareg' to fit our models (Cribari-Neto and Zeileis 2010).

interval, such as ratios or proportions (see Appendix A4 for details on the distribution of our ratio variable).

Three more issues concerning our data and model need to be mentioned. First, Beta Regression only allows for values between 0 and 1. As our ratio variable also assumes the lower extreme of 0, we use the transformation proposed by Cribari-Neto and Zeileis (2010: 3)

$$\hat{y} = \frac{y \times (n-1) + 0.5)}{n}$$

Where *y* is our dependent variable and *n* is the sample size. Second, our model accounts for the length of IO treaties, that is, the absolute number sentences, and thus for the maximum of hits our dictionary could have found by setting the *weights* function to the total number of sentences in a given treaty. Third, collinearity is not an issue in our analysis as a variance inflation factor (VIF) test shows (see Appendix A3). None of the variables' VIF exceeds a value of two (the conservative threshold is five). Descriptive statistics are presented in Table 2.

Table 2: Descriptive Statistics

	Observations	Mean	Std. Deviation	Minimum	Maximum
Incomplete Ratio	72	0.08	0.04	0	0.21
Delegation	72	0.14	0.14	0	.56
Pooling	72	0.25	0.19	0	.73
Membership	72	24.49	27.71	3	134

We estimate a Beta regression model, as shown in Table 3. The model estimates the effects for the main body of the treaty. To facilitate the interpretation of the multivariate analysis, we z-standardize all continuous explanatory variables.

Table 3: Beta Regression of incompleteness

	Estimate	Standard Error	Exp. Coefficients	95% Conf. Interval
Intercept	-2.26***	0.004	0.105	0.104, 0.105
Delegation	0.14***	0.003	1.152	1.146, 1.159
Pooling	-0.05***	0.004	0.949	0.041, 0.957
Members	0.05***	0.004	1.056	1.048, 1.063
Observations	72			
Log Likelihood	33,410.0			
Pseudo R^2	0.05			
Phi	62.12			
Note:	*p<0.05; **p<0.01; ***<0.001			

Note: Exponentiated coefficients represent odds ratios.

Given the hard to interpret nature Beta regression coefficients, Table 3 also presents the results as exponentiated coefficients. Each exponentiated coefficient represents a change in the ratio of proportions, given a one-unit change in the explanatory variable. To illustrate the interpretation of the results from Table 3, consider the effect of a one-unit change in the *members* variable on the share of incomplete sentences. Look at the *members* row of exponentiated coefficients. All else equal, a one-unit change in this variable, that is, an increase of one standard deviation (27.71) leads to a relative change of 1.056 in the share of incomplete sentences. In other words, for an IO with 28 more members, the share of incomplete sentences increases by five percent.

The model shows that our three explanatory variables have significant statistical effects on incompleteness. Both the level of delegated authority and the number of member states lead to more incomplete contracts. By contrast, higher levels of pooled authority lead to less incomplete contracts. These results support our three hypotheses. As expected, states tend to design more incomplete IO treaties when the underlying structure of the cooperation problem requires them to delegate more authority to international bureaucracies. All else equal, a one-unit increase in the level of delegated authority (*SD* = 0.14), results in a 15 percent higher share of incomplete sentences. This has the consequence that IOs with a high level of delegated authority are additionally equipped with ample opportunities to interpret their mandates. These IOs are thus not only formally empowered to perform a broad array of tasks in different issue areas, but they also have broad discretion interpreting and performing these delegated tasks. Prominent examples of this constellation include the EU, the European Economic Area (EEA), and ILO. But less prominent organizations, such as the Andean Community (CAN) are also part of this group of IOs. Similarly, the incompleteness of IO treaties tends to increase with the number of founding member states. All else equal, a one-unit increase in the number of founding members (*SD* = 27.71), results in a five percent higher share of incomplete sentences. As expected, larger groups of states, therefore, seem to have greater difficulties in agreeing on precise treaty language.

Finally, we also find a significant statistical effect for pooled authority. IOs in which states take collective decisions without being able to (unilaterally) block them beget more precise treaty language to minimize this prospect. All else equal, a one-unit increase in the level of pooled authority (*SD* = 0.19), results in a five percent lower share of incomplete sentences. Our sample of IOs includes a number of prominent examples for this constellation, like the ISA, the Bank for International Settlements (BIS) or the World Health Organization (WHO). Especially the case of the WHO might point to the potential consequences of a high level of pooled authority under precise contracts. For decades, consecutive Directors-General

have tried to reform the organization, making it fit for new challenges and a growing number of competing actors in the field of global health. Yet, these reform processes failed repeatedly (Hanrieder 2015). A potential reason for this inertia might be that states have to agree on collective decisions under a very precise set of rules. Fritz Scharpf (2006) had once alluded to this phenomenon under the label "joint decision trap."

Robustness checks

To evaluate the robustness of our results, we estimate a set of alternative models (see Appendix A6). First, we use the count of incomplete sentences as our main dependent variable and use a negative binomial regression model with the offset function set to the total number of sentences in a given treaty. This presupposes that 10 incomplete terms in a treaty of 100 words length are less open to subsequent interpretation than a treaty of 1000 words length containing 100 incomplete terms. Similar to the Beta regression model for the share of incomplete sentences, this model yields almost identical results. While delegation and membership have statistically highly significant positive effects of the count of incomplete sentences, pooling has a statistically highly significant negative effect.

Second, as logistic regression models are often suggested for analyzing proportional data as our ratio variable, we also estimate a logistic regression model to analyze the effect of our explanatory variables on the share of incomplete sentences. With the exception of slightly lower levels of statistical significance for pooling and members, the results of this model are almost identical to those presented in Table 3. Overall, this shows that our results are robust to different specifications of our dependent variable and statistical model.

V. Conclusion

In this paper, we have contributed to the literature on the rational design of international institutions by explaining IO constitutions from an incomplete contract theory perspective. We submit that an important facet of incompleteness manifests itself in ambiguous or vague language, leaving ample scope for interpretation among actors in the aftermath of agreements being first agreed. Our results show that incompleteness is shaped by delegation, pooling, and the number of founding states. IOs with secretariats that have been transferred greater decision-making authority show higher levels of incompleteness. We interpret this as states being willing to cede more ground through flexible language in areas where they agreed that setting up a powerful international secretariat to solve collective action problem is in their best interest. Moreover, the number of member states involved in creation of an IO increases incompleteness. The causal mechanism at work here is the need for multiple actors to agree on a common treaty text, which is aided by ambiguity and the ability to interpret the text in more than one way. Finally, pooling reduces incompleteness, as we hypothesized, because of the sovereignty costs involved in this decision. Given the nature of our research, we do not make strong causal claims, even though we believe that the causal mechanisms spelled out are well corroborated by previous research. Nevertheless, future research could take the results from our study and assess whether the causal mechanism we posit and correlations we found also hold up in the context of in-depth, qualitative case studies.

As far as our automated approach to measure incompleteness is concerned, we see ample scope for application in a variety of different contexts such as acts, regulations, and international agreements, more broadly. The approach developed in this paper, therefore, spells implications beyond political science. We consider our approach suitable for all texts where a fair amount of effort has gone into their preparation and that hinge on compromises

among independent actors. If these scope conditions are met, our approach can be used in political science just as much as economics or law. On the other hand, other classes of documents such as political speeches, transcripts of hearings, or party manifestos carry little promise of yielding substantively meaningful results. However, since this class of documents is unlikely to be approached from an incomplete-contracts perspective in the first place, we consider this limitation natural rather than problematic. Consequently, there is vast scope for future applications of our approach, which could become an interesting alternative operationalization of incompleteness whenever machine-readable documents can be obtained. This will open new avenues for future research in this area and, over time, refine our understanding of how incompleteness can be accounted for theoretically.

References

Abbott, K. W., Keohane, R. O., Moravcsik, A., Slaughter, A.-M. and Snidal, D. (2000) 'The Concept of Legalization', *International Organization* 54(3): 401–419.

Abbott, K. W. and Snidal, D. (2000) 'Hard and Soft Law in International Governance', *International Organization* 54(3): 421–456.

Abbott, K. W. and Snidal, D. (2010) 'International regulation without international government: Improving IO performance through orchestration', *The Review of International Organizations* 5(3): 315–344.

Barker, C. (2002) 'The Dynamics of Vagueness', *Linguistics and Philosophy* 25(1): 1–36.

Barker, C. (2010) 'Vagueness', in K. Allan (ed.). *Concise Encyclopedia of Semantics*. Amsterdam: Elsevier, pp. 1037–1040.

Barnett, M. and Finnemore, M. (2004) *Rules for the World: International Organizations in Global Politics*, Ithaca, NY: Cornell University Press.

Best, J. (2005) *The Limits of Transparency: Ambiguity and the History of International Finance*, Ithaca, NY: Cornell University Press.

Best, J. (2012) 'Ambiguity and Uncertainty in International Organizations: A History of Debating IMF Conditionality', *International Studies Quarterly* 56(4): 674–688.

Blake, D. J. and Payton, A. L. (2015) 'Balancing Design Objectives: Analyzing New Data on Voting Rules in Intergovernmental Organizations', *The Review of International Organizations* 10(3): 377–402.

Chayes, A. and Chayes, A. H. (1993) 'On compliance', *International Organization* 47(2): 175–205.

Cooley, A. and Spruyt, H. (2009) *Contracting States: Sovereign Transfers in International Relations*, Princeton, NJ: Princeton University Press.

Cribari-Neto, F. and Zeileis, A. (2010) 'Beta Regression in R', *Journal of Statistical Software* 34(1): 1–24.

Däubler, T., Benoit, K., Mikhaylov, S. and Laver, M. (2012) 'Natural Sentences as Valid Units for Coded Political Texts', *British Journal of Political Science* 42(4): 937–951.

Dür, A. and Lechner, L. (2019) *Humans versus computers: Manual and computational estimates of the design of international institutions*, available at http://tradepower.sbg.ac.at/docs/human_vs_computer_preprint.pdf (accessed January 2019).

Ehrlich, I. and Posner, R. A. (1974) 'An Economic Analysis of Legal Rulemaking', *The Journal of Legal Studies* 3(1): 257–286.

Endicott, T. A. O. (2000) *Vagueness in Law*, Oxford: Oxford University Press.

Epstein, D. and O'Halloran, S. (1999) *Delegating Powers: A Transaction Cost Politics Approach to Policy Making Under Separate Powers*, Cambridge: Cambridge University Press.

Farrell, H. and Héritier, A. (2007) 'Introduction: Contested competences in the European Union', *West European Politics* 30(2): 227–243.

Ferrari, S. and Cribari-Neto, F. (2004) 'Beta Regression for Modelling Rates and Proportions', *Journal of Applied Statistics* 31(7): 799–815.

Garrett, G. (1992) 'International cooperation and institutional choice: the European Community's internal market', *International Organization* 46(2): 533–560.

Goldstein, J., Kahler, M., Keohane, R. O. and Slaughter, A.-M. (2000) 'Introduction: Legalization and World Politics', *International Organization* 54(3): 385–399.

Grimmer, J. and Stewart, B. M. (2013) 'Text as Data: The Promise and Pitfalls of Automatic Content Analysis Methods for Political Texts', *Political Analysis* 21(3): 267–297.

Hanrieder, T. (2015) *International Organizations in Time: Fragmentation and Reform*, Oxford: Oxford University Press.

Hart, O. and Holmström, B. (1987) 'The Theory of Contracts', in T. F. Bewley (ed.). *Advances in Economic Theory: Fifth World Congress*. Cambridge, MA: Cambridge University Press, pp. 71–155.

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

Helleiner, E. (2014) *Forgotten Foundations of Bretton Woods: International Development and the Making of the Postwar Order*, Ithaca, NY: Cornell University Press.

Héritier, A. (2012) 'Institutional Change in Europe: Co-decision and Comitology Transformed', *Journal of Common Market Studies* 50: 38–54.

Hooghe, L. et al. (2017) *Measuring International Authority: A Postfunctionalist Theory of Governance, Volume III*, Oxford: Oxford University Press.

Hooghe, L. and Marks, G. (2009) 'A postfunctionalist theory of European integration: from permissive consensus to constraining dissensus', *British Journal of Political Science* 39(1): 1–23.

Huber, J. D., Shipan, C. R. and Pfahler, M. (2001) 'Legislatures and Statutory Control of Bureaucracy', *American Journal of Political Science* 45(2): 330–345.

Keefe, R. (2000) *Theories of Vagueness*, Cambridge: Cambridge University Press.

Koremenos, B. (2013) 'What's left out and why? Informal provisions in formal international law', *The Review of International Organizations* 8(2): 137–162.

Koremenos, B. (2016) *The Continent of International Law: Explaining Agreement Design*, Cambridge: Cambridge University Press.

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

Lake, D. A. (2007) ‘Delegating divisible sovereignty: Sweeping a conceptual minefield’, *The Review of International Organizations* 2(3): 219–237.

Lenz, T., Bezuijen, J., Hooghe, L. and Marks, G. (2014) ‘Patterns of International Organization: Task Specific vs. General Purpose’, *Politische Vierteljahresschrift* 49: 131–156.

Linos, K. and Pegram, T. (2016) ‘The Language of Compromise in International Agreements’, *International Organization* 70(3): 587–621.

Manger, M. S. and Peinhardt, C. (2017) ‘Learning and the Precision of International Investment Agreements’, *International Interactions* 1–21.

Mattli, W. and Stone Sweet, A. (2012) ‘Regional Integration and the Evolution of the European Polity: On the Fiftieth Anniversary of the Journal of Common Market Studies’, *Journal of Common Market Studies* 50: 1–17.

Mellinkoff, D. (1963) *The Language of the Law*, Boston, MA: Little, Brown.

Moravcsik, A. (1993) ‘Preferences and power in the European Community: a liberal intergovernmentalist approach’, *Journal of Common Market Studies* 31(4): 473–524.

Moravcsik, A. (1998) *The Choice for Europe: Social Purpose and State Power from Messina to Maastricht*, Ithaca, NY: Cornell University Press.

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

Pevehouse, J. C., Nordstrom, T. and Warnke, K. (2004) ‘The Correlates of War 2 International Governmental Organizations Data Version 2.0’, *Conflict Management and Peace Science* 21(2): 101–119.

Pollack, M. A. (1997) ‘Delegation, Agency, and Agenda Setting in the European Community’, *International Organization* 51(1): 99–134.

Scharpf, F. W. (2006) ‘The Joint-Decision Trap Revisited’, *Journal of Common Market Studies* 44(4): 845–864.

Schmidt, S. K. (2018) *The European Court of Justice and the Policy Process: The Shadow of Case Law*, Oxford: Oxford University Press.

Tallberg, J. (2002) 'Delegation to Supranational Institutions: Why, How, and with What Consequences?', *West European Politics* 25(1): 23–46.

Thatcher, M. and Stone Sweet, A. (2002) 'Theory and Practice of Delegation to Non-Majoritarian Institutions', *West European Politics* 25(1): 1–22.

Tirole, J. (1999) 'Incomplete Contracts: Where do We Stand?', *Econometrica* 67(4): 741–781.

United Nations (2018) *United Nations Treaty Collection*, available at www.treaties.un.org.

Volgy, T. J., Fausett, E., Grant, K. A. and Rodgers, S. (2008) 'Identifying Formal Intergovernmental Organizations', *Journal of Peace Research* 45(6): 837–850.

Williamson, O. E. (2000) 'The New Institutional Economics: Taking Stock, Looking Ahead', *Journal of Economic Literature* 38(3): 595–613.