Boilerplate in International Trade Agreements*

Claire Peacock
École supérieure d'études internationales, Université Laval and
Department of Political Science, Simon Fraser University
claire.peacock.1@ulaval.ca

Karolina Milewicz
Department of Politics & International Relations and University College
University of Oxford
karolina.milewicz@politics.ox.ac.uk

Duncan Snidal
Department of Politics & International Relations and Nuffield College
University of Oxford
duncan.snidal@politics.ox.ac.uk

Abstract

New international agreements often recycle language from previous agreements, using boilerplate solutions alongside customized provisions. The presence of boilerplate in international agreements has important implications for understanding how international rules are made. The determinants behind boilerplate in international agreements have not previously been systematically evaluated. Using original data from a sample of 348 preferential trade agreements (PTAs) adopted between 1989 and 2009, we combine novel text analysis measures with Latent Order Logistic Graph (LOLOG) network techniques to assess the determinants behind boilerplate in labor and environmental provisions commonly found in PTAs. Our results indicate that whereas boilerplate can be used for both efficiency and distributive purposes, international boilerplate is used primarily for efficiency gains and power-distribution considerations are not systematically important.

*For helpful suggestions on previous drafts of this paper we thank the anonymous reviewers, Jeff Checkel, Andy Eggers, Aaron Hoffman, Alexander Kentikelenis, Larry King, Bernhard Reinsberg, Oskar Timo Thoms, Laurel Weldon, and other participants at the University of Cambridge International Political Economy research seminar 2017 and at the International Studies Colloquium at Simon Fraser University 2018. For research assistance, we thank Jackie Majnemer. For updating the LOLOG package to facilitate our analysis, we thank Ian Fellows and Mark Handcock. We also acknowledge generous support from the Swiss National Science Foundation.

By clicking "I agree" to the terms and conditions associated with online purchases or other internet-based services, individuals engage in everyday forms of contracting. As most customers know, the fine print found in these contracts generally relies on "boilerplate" – standardized language that is used across multiple agreements (Smith 2006, 1176). Boilerplate is a common tool of contract design, observed in all forms of law, but its use is best understood in the context of domestic contracts.[1] Boilerplate provides a way to lower the costs of contracting, reduce uncertainty (Kahan and Klausner 1997), increase efficiency (Gulati and Scott 2012) and incorporate "accumulated wisdom" (Hill 2001, 59). Yet, relying on boilerplate also has drawbacks; it can perpetuate familiar but inadequate solutions at the expense of innovation in the drafting of contracts to suit different circumstances (Hill 2001). Moreover, boilerplate is often provided by one party – potentially at the expense of another – so that it reflects asymmetric power rather than the ideal of a jointly negotiated contract.[2]

The use of boilerplate in international treaties "concluded between States in written form and governed by international law" (Article 2(1), Vienna Convention on the Law of Treaties 1969) is less well understood. Despite the proliferation of international relations research on legalization (Goldstein et al. 2000) and treaty design (Koremenos 2013) that seeks to explain treaty content (Böhmelt and Spilker 2016; Hooghe and Marks 2015) and related outcomes (Hafner-Burton 2005), much less consideration has been given to the legal wording of these agreements, which can be custom-made, boilerplate, or a combination.[3]

Boilerplate in international agreements as a form of contract[4] has started to draw attention, especially in the field of international political economy. New bodies of work on bilateral investment treaties (Alschner and Skougarevskiy 2016; Brown 2013; Poulsen 2014; Vandevelde 1992), the diffusion of regional institutions (Alter 2012; Arnold and Rittberger 2013; Börzel and Risse 2012; Jetschke and Murray 2012; Lenz 2012) and the transplantation of laws (Morin and Gold 2014) acknowledge the growing use of legal models and templates in international relations. Similarly, recent scholarship on the design of trade agreements (Allee and Elsig 2016; Allee, Elsig and Lugg 2017; Alschner, Seiermann and Skougarevskiy 2018; Baccini, Dür and Haftel 2015; Morin, Pauwelyn and Hollway 2017) observes boilerplate in economically important and politically sensitive preferential trade agreements (PTAs), also with regard to provisions on dispute settlement mecha-

[1]Boilerplate is used in legal literature to describe "standard-form contracts" not subject to negotiation. It contrasts with "law of negotiations" which is the result of a bargain. See Ben-Shahar (2005/2006, 821).

[2]On the one-sidedness of boilerplate contracts see Bebchuk and Posner (2005/2006, 827).

[3]For an exception see Linos and Pegram (2016).

[4]Though some differences exist, international treaties are often understood as a form of contract. Oppenheim (1955, 877) famously argues that "[i]nternational treaties are agreements, of a contractual character, between States, or organisations of States, creating legal rights and obligations between the Parties". Baker (2005, 1324) suggests that "free trade agreements can be better understood as business contracts, building a common body of law to be used by the partners to the contract".

nisms (Jo and Namgung 2012). In EU PTAs, for example, the now standard "sustainable development" chapters contain nearly identical language based on the NAFTA model. The US also relies on boilerplate in its trade agreements. The 2006 Peru-US Free Trade Agreement (FTA) and the 2007 Panama-US FTA, for example, use identical wording (including capitalization) in their preambles regarding the need to "PROTECT, enhance, and enforce basic workers' rights, [and] strengthen their cooperation on labor matters".

This previous scholarship regarding standard-form international contracting has been primarily descriptive in revealing patterns of PTA text similarity. The reasons behind states' use of boilerplate in international agreements remain largely unexplained. Our analysis is thus guided by the following question: "when and why do states use boilerplate in their international agreements?" We theorize two broad rationales for boilerplating in international agreements. First, boilerplate in international agreements may be chosen for its *efficiency-related benefits*. Compared to tailored solutions, boilerplate minimizes costs related to the preparation of a legal text, negotiation, implementation, and compliance. Second, boilerplating may be adopted because of its *distributional effects*, privileging powerful states that impose it on their weaker contracting partners.[5] More specifically, powerful states may use boilerplate to spread their preferred regulation or other legal solutions through their network of treaty partners. While we have framed these as primarily strategic choices, they may equally be the result of boilerplate spreading through parallel social processes such as mimicking existing agreements (Goodman and Jinks 2013; Johnston 2007) and norm diffusion (Acharya 2013; Keck and Sikkink 2004), as we discuss below.

We contribute to the understanding of boilerplate in international agreements by investigating these two main rationales behind the use of boilerplate in (main text) PTA provisions addressing non-trade issues (NTIs) which Milewicz et al. (2018, 744) point out "do not directly concern trade and go beyond what is typically regulated by the multilateral trading system". Because PTAs among states are common, they provide a rich and sufficiently large data set in which to explore boilerplating. Trade agreements frequently include NTI provisions promoting environmental standards or regulating domestic labor conditions that bear no direct or significant relation to the central trade-promoting purposes of PTAs. These NTIs – while appearing in agreements that are binding upon their state parties and which are therefore considered hard law – are nonetheless written in general and imprecise language more typical of soft law. As soft law is more likely than hard law to disseminate through agreements (Böhmelt and Spilker 2016), NTIs provide an excellent location for detecting boilerplate and exploring the different motives behind its use.

In order to assess these effects empirically, we employ a novel combination of text-similarity methods and statistical network analysis. By network, we mean a system of

[5]On powerful states using international law to codify their preferences see Drezner (2001; 2007).

interconnected objects, here agreements, which are linked insofar as they share boilerplate. We are trying to explain the observed pattern of interconnection. Our analysis is based on a sample of NTI provisions in 348 English-language[6] PTAs adopted between 1989 and 2009, of which 107 include environmental provisions and 135 include labor rights provisions.[7] Using machine-readable versions of the PTA texts, we compute cosine similarity scores between all pairs of PTA texts – based on sequences of five words (i.e. 5-grams) – from which we construct text-similarity networks of their NTI clauses. Each tie in the resulting networks represents the presence of shared boilerplate between a pair of agreements. We then employ Latent Order Logistic Network Models (LOLOG) (Fellows 2018), a new statistical network analysis adapted for evaluating growth networks, to test our hypotheses.

Our results indicate that NTI boilerplate is first and foremost driven by efficiency-related considerations. Boilerplate offers clear efficiency benefits, but it does not reflect distributive aspects, i.e. the adaptation of international agreements to meet the needs of more powerful states. Our findings are robust to a series of alternate model specifications, with only slight variations. These include the use of an alternate text similarity measure, varied word sequence length, and the inclusion of additional control variables, as presented in Appendix C (pages 8–10).

1 What is boilerplate?

The term "boilerplate" describes a "standard form of words used in drafting contracts or other legal documents" (Law and Martin 2014). In practice, boilerplate is identified by the presence of nearly identical language across clauses or agreements. For example, Articles II(1) in both the bilateral investment treaties between Canada and Soviet Union (1989) and between China and Russia (2006) begins: "Each Contracting Party shall aspire to create favorable conditions to investors of the other Contracting Party to make investments in its territory" (Ben-Shahar 2005/2006, 821).[8]

Boilerplate at the international level differs from standard domestic boilerplating in one important respect. In the domestic setting, contracts (and boilerplated clauses) are typically "hard law" and can be enforced. International agreements, however, sometimes entail "soft" elements that cannot be legally enforced, although they put normative, reputational and other pressures on states (Abbott and Snidal 2000). States may include

[6]It would be difficult to accurately detect text re-use across languages. We were unable to locate English-language versions of approximately 70 PTAs signed between 1945 and 2009, primarily from Latin America and the Middle East. None of these excluded PTAs had labor provisions and only four included environmental provisions.

[7]For details see Design of Trade Agreements (DESTA) Database (2018).

[8]Close variations of this boilerplate also appear in other bilateral investment treaties, e.g. Article 3(1) of the treaty between the United Kingdom and India (1994). On the use of boilerplate in bilateral investment treaties see Poulsen (2015).

such clauses in order to pay symbolic lip service, even if they do not want to see them seriously enforced (Peacock 2018). Indeed, NTIs are often written in soft or imprecise language such as "shall aspire to create favorable conditions". Thus, boilerplate may be used either to achieve certain ends or to acknowledge ends without necessarily achieving them (Spilker and Böhmelt 2013).

Although the term "boilerplate" is sometimes used interchangeably with "template" or "model text", it is conceptually distinct. Whereas "boilerplate" denotes the use of specific wording and thus the close similarity of both content and language, "template" designates only a general similarity of content and not the specific wording of provisions. A template is a guide, often used to indicate that a certain type of provision, such as a clause on child labor, should be included in every agreement signed by a particular party. The US Bipartisan Trade Promotion Authority Act of 2002, for example, required each US FTA to address NTI-related objectives such as ensuring "that trade and environmental policies are mutually supportive" (U.S. Congress 2001/2002) but did not provide exact language. "Model" is a yet broader term that encompasses both boilerplate and templates. For example, states often publish model bilateral investment treaties – a practice which is uncommon in the world of PTAs – from which they sometimes draw specific boilerplate language and other times only a rough guide for what to include in their agreement (Brown 2013).

Thus, while boilerplate, template and model all address the similarity of substance across legal agreements, they are cast at different levels of detail. Model is usually at the most general level, template is focused on the overall similarity of legal content, and boilerplate is most specific in matching exact legal wording across agreements. We concentrate on boilerplate as the most stringent requirement, and also one that can be operationalized clearly (see section 4.2), therefore providing the strongest and clearest evidence of cross-agreement borrowing.

2 Why do states use boilerplate in international agreements?

When drafting international agreements, states must both choose their design features (Koremenos, Lipson and Snidal 2001) – such as scope, membership, and flexibility – and also decide how to express these choices in terms of language, i.e. whether through customized text or by relying on existing boilerplate. The choice of wording in international agreements reflects the underlying motives behind the inclusion of a particular feature and has important implications for how a specific provision impacts state behavior (Linos and Pegram 2016). Thus, the specific language used in an agreement matters and is a worthy subject of study. In this article we pursue the question of why states use boilerplate in international agreements and explore two alternative rationales: efficiency-related reasons and distributional goals adapted to the circumstances of particular issue areas.

2.1 Efficiency

Boilerplating increases efficiency by allowing states to reduce various costs associated with making an international agreement. First, using standard clauses reduces the costs of drafting an agreement and takes fewer legal resources to draft and "scrub" than do tailor-made treaty texts. Boilerplate provides a focal solution that avoids the costs of niggling over a myriad of small details that each require specialized legal teams.[9] Thus, the familiarity of boilerplate text may facilitate agreement and reduce negotiating costs.

Second, boilerplate may be deemed "appropriate" because it has been consistently used, especially by more experienced states that have demonstrated appropriate behavior in earlier agreements (Checkel 2005; March and Olsen 1998). In the same vein, "mimicking is a microprocess whereby a novice copies the behavioral norms of the group ...[as] an efficient means of adapting to uncertainty prior to any detailed ends-means calculations of the benefits of doing so" (Johnston 2007, 23).[10] In effect, both states and negotiators may be socialized to adopt familiar solutions, which leads to efficiency gains.

Third, boilerplate lowers costs of implementation and compliance insofar as it creates uniformity among clauses across different agreements for participating states (Kelley 2008) enabling a more streamlined implementation and compliance process. When used repeatedly by a state, boilerplate may ensure that a given text is consistent with a state's domestic legislation, reducing the costs associated with compatibility of new laws.

Fourth, and related, boilerplate can reduce uncertainty about the likely legal interpretation and impact of an agreement as states can draw on past experience with familiar clauses to better anticipate their consequences. In short, borrowing from prior contracts is a means to "learn quickly and efficiently from their experience and from that of others" (Hill 2001, 60) and to promote best practices learned through trial and error. Accordingly, many of these arguments involve positive network externalities: as more states and NTIs use boilerplate, it becomes increasingly efficient for subsequent states and agreements to use the same boilerplate.

Insofar as achieving efficiency in international agreements drives boilerplating, we should observe an increasing tendency for states to copy clauses from past agreements

[9] Gulati and Scott (2012, 6) describe boilerplate in sovereign debt contracts as a "three and a half minute" solution for lawyers facing a trade-off between spending time on drafting a new contract "and the [much lower] effort costs of redesigning boilerplate that was widely used and had been part of the standard form contract for many years".

[10] In diffusion analyses, this mechanism is also referred to as "learning", whereby policy makers are inclined to adopt policies that have proven successful elsewhere (Gilardi 2010). For related analyses of international policy diffusion as a consequence of economic interdependence see e.g. (Elkins, Guzman and Simmons 2006; Simmons, Dobbin and Garrett 2006; Simmons and Elkins 2004).

into future ones. In terms of network analysis, there are two related ways to detect efficiency-related mechanisms in boilerplating. One is that as particular clauses become popular, we should see them appearing increasingly frequently in future NTI clauses. In network terms, this is called "degree popularity", which is explained below. While this may occur simply through general imitation and diffusion of earlier contractual language, it also leads to efficiency gains. A second mechanism is through direct contact involving a "common member" promoting the dissemination of an NTI clause that it used in its own prior agreements.

2.2 Power and distribution

Our discussion of the efficiency gains of boilerplating assumes that the issues are not highly contentious and that states share largely common interests and values in designing their agreements. Of course, many international agreements involve significant differences among states and entail potentially serious bargaining over their exact terms. We capture those considerations in terms of distributive differences.

According to the distribution-related explanation, boilerplate is a means of exerting power over legal design to improve a party's bargaining position (Schwartz and Scott 2003)), to spread its preferred rules (Allee and Elsig 2016) or to promote its normative ideals (Goodman and Jinks 2013). Regardless of the motivation, boilerplate can make certain clauses seem appropriate and non-negotiable (Checkel 2005; March and Olsen 1998). As one author puts it: "Even in negotiated contracts, lawyers may use the *boilerplate* label as a strategic ploy to signal that the terms are not negotiable or are so standard that only a fool would question them" (Clapp et al. 2011). Thus, boilerplating has the effect of influencing trade partners to accept another state's preferred legal solutions or to risk derailing negotiations altogether.[11] For example, when the United Kingdom was seeking to secure post-Brexit trade deals, prospective trading partners demanded it back off from using the boilerplate human rights standards imposed in EU trade agreements (Partington 2019).

We hypothesize that distributional considerations will lead influential states to include boilerplate language in the design of their PTAs. High income states have the greatest ability to achieve their desired legal outcome in negotiations as their substantial legal expertise combined with their economic capacity enables them to influence reluctant partners into accepting their preferred language. In our network analysis, we capture these powerful states' likelihood of using boilerplate that includes their preferred legal language by evaluating whether a pair of NTIs both contain "high income" states as members. In addition, we combine our measure of "high income" with "common member" (to

[11] Distributive considerations are similarly reflected in "conditionality" requirements imposed on borrower states by international financial institutions, such as the World Bank and IMF (Mosley, Harrigan and Toye 1995; Schimmelfennig 2008; Vreeland 2003).

"high-income common member") to depict potential conditionality of efficiency-related reasons for boilerplating on distributional considerations.

The use of boilerplate may also depend on exactly how the substance of the provisions under consideration shapes distributive considerations. Boilerplating NTI clauses represents an effective way for powerful states to exert their influence over the content of the agreement without bearing the costs of legal "tailoring". Where powerful states themselves follow high standards on an issue, they may require that PTAs incorporate NTI provisions that reflect these standards. This argument is analogous to the "California effect" (Vogel 1995; 1997; Perkins and Neumayer 2010), whereby access to markets with stricter standards is made conditional on other states' adopting equivalent high standards to "level the playing field". Conversely, if powerful states do not follow high standards, they may find that boilerplate provides a convenient form of "window dressing" that makes it appear that attention is being paid to such standards without actually imposing costs or requirements on participating states. This is also compatible with Checkel 's (2005) "role playing" mechanism where actors comply with group norms in a nonreflective manner and without necessarily paying attention to consequences.

3 Identifying boilerplate in NTIs and PTAs: Data and methods

Our analysis focuses on boilerplating with regard to two NTIs addressed in PTAs – labor rights and environmental clauses. It is based on original data gathered from a sample of 348 English-language PTA texts adopted between 1989 and 2009, 135 agreements including labor provisions and 107 environmental provisions.[12] Most of these agreements are between states but some also include groupings such as EFTA or the EU. We use the terms "states" to refer to both types of actors for simplicity. Boilerplate "ties" between agreements are defined in terms of high levels of sharing common text.

To assess the determinants of boilerplate in NTI clauses, we pair a standard text similarity measure – cosine similarity[13] – with the Latent Order Logistic Graph (LOLOG) statistical network modelling technique. LOLOG models are appropriate for studying growth networks where ties are formed but are not subsequently broken, such as in our NTI networks. Our analysis represents the first application of LOLOG modeling to political science data.

[12]While our PTA sample covers the entire post-World War II period, a natural starting point for our analysis is 1989 when PTAs first began to proliferate and when treaty texts became more widely available.

[13]For other uses of text-analysis tools in political science see Grimmer and Stewart (2013).

3.1 Data processing

In order to investigate boilerplating across PTAs, we converted the text of each PTA into a machine-readable file. From these files, we extracted the text of each NTI clause and created separate sub-files for all labor and environmental provisions. The resulting files of NTI texts then underwent standard text-preparation steps for natural language processing: "word-stemming"; removing extraneous characters such as page numbers, symbols, and punctuation; and changing all characters to lowercase to ensure the raw texts were comparable.[14] In addition, we removed "stop words", such as country names and nationalities, used to adapt boilerplate to specific conditions.[15]

We then broke our texts down into sequences of words and counted their respective frequencies within each text. While many text analyses use the standard "bag-of-words" approach according to which counts of single words are assessed as if drawn from a bag of words in a random order (e.g. Burscher et al. 2014; Hopkins and King 2010; Sebastiani 2002), such an approach is by definition not sensitive to the context or order in which the words appear - something that is the essence of detecting boilerplate. To capture this, we therefore analyzed sequences of five words (5-grams)[16] – a number chosen to balance between the context of words and the possibility to detect partial copies[17] across two texts that have different words surrounding them.

To demonstrate the concept of a 5-gram, consider the following sentence from Article 1.2 of the 2007 United States-Panama Trade Promotion Agreement:

> "The objectives of this Agreement, as elaborated more specifically through its principles and rules, including national treatment, most-favored-nation treatment, and transparency, are to:"

Ignoring punctuation and capitalization, the first 5-gram of this sentence is "the objectives of this agreement"; the second 5-gram in turn is "objectives of this agreement as"; the third is "of this agreement as elaborated", and so on.

[14] These steps were performed in R, Quanteda library, version 0.9.9-65 (Benoit 2017). Stemming involves the removal of word endings to leave a common base form of each word. For example, "democracy", "democratization", and "democratic" are all reduced to "democra" as their stemmed form. To do this we used the standard Porter stemmer implemented within Quanteda.

[15] The term "stop word" refers to any word that is removed from a text before or after it is processed. Our custom list of stop words is provided in Appendix B (p. 5). In standard contracts, stop words correspond to the blank spaces that must be filled in by the parties to the contract (e.g., their names).

[16] An "n-gram" is a sequence of terms that appears in a text. On the utility of the "n-gram" approach see Damashek (1995). A 5-gram specification has also been used in a previous study of PTA text similarity (Alschner, Seiermann and Skougarevskiy 2018). As a robustness check, we also calculated 3-grams (see Appendix C, p. 8). Longer sequences of words (e.g. 7-grams) dramatically limit the ability to detect partial copies and are therefore not used here.

[17] "Partial copy" refers to a text that is not an identical copy of another text, but that shares some of its terms. Boilerplate, which often requires small changes to adapt it to specific contexts, such as updating names or places, is usually a partial copy rather than an exact copy of another text.

3.2 Similarity measure

To analyze each distinct NTI 5-gram, we created vectors of 5-gram counts. The resulting vectors – a numerical representation of our texts – were then used to calculate cosine similarity scores for each possible pairing of NTI clauses, our main measure of boilerplate.[18] The cosine similarity score measures the angle between two vectors of counts. If the two vectors are identical, the angle is 0 degrees and its cosine is 1. A similarity score of 1 indicates that all 5-grams found in a pair of PTA texts are the same; it is a clear marker of boilerplate. If the vectors share no features, the resulting angle is 90 degrees and its cosine is 0. In this case, there is no boilerplate. If some boilerplate is shared, the cosine value will be between 0 and 1, with higher values representing a greater amount of shared boilerplate. To calculate cosine similarity, $\cos\theta$, we use the following formula:

$$\text{similarity}(A,B) = \cos\theta = \frac{\sum_{i=1}^{n} A_i B_i}{\sqrt{\sum_{i=1}^{n} A_i^2}\sqrt{\sum_{i=1}^{n} B_i^2}}$$

where A_i and B_i are each vectors of counts and n is the length of the vectors. The cosine similarity score returns the number of common counts divided by the number of possible counts. It is comparable to a correlation coefficient, though cosine similarity is more commonly applied to assess text similarity, such as in plagiarism detection software.[19] In general, higher cosine similarity scores indicate an increased presence of boilerplate.

Cosine similarity has particularly useful properties for text analysis. First, it is length invariant. Unlike alternatives, such as the Jaro-Winkler distance, which are only applicable to texts of equal length, cosine similarity can be calculated for texts of different lengths. Second, cosine similarity is invariant to repetition. This means that if a sequence of words (i.e., 5-gram) appears only once in one text but multiple times in the other, the difference in frequency of occurrences does not artificially deflate or otherwise affect the resulting score. The extended Jaccard similarity measure, for instance, is sensitive to such repetition. Third, applied to texts, cosine similarity scores assume values ranging from 0 (indicating no relation between two sequences of words) to 1 (indicating a perfect match). In contrast to measures that are not bounded on [0,1] – such as the Jaro-Winkler or Levenshtein distances – this enables a straightforward interpretation of statistics across pairs of texts.

To illustrate the properties of the cosine similarity score, consider the following two clauses regarding the recognition of labor rights as provided by International Labour

[18] Less common alternatives include: extended Jaccard similarity; the Jensen-Shannon divergence (for application see Hollway and Koskinen 2016); distance-based measures, such as the Jaro-Winkler distance or Levenshtein distance. For an overview of similarity and distance measures see Tan, Steinbach and Kumar (2007).

[19] Cosine similarity and Pearson correlation are both based on the inner product of two vectors and are bounded on [-1,1]. If applied to texts where there are no negative values, $\cos(x,y)$ is bounded on [0,1]. Correlation is the cosine similarity between centered vectors and is equivalent to $\cos(x-\bar{x}, y-\bar{y})$ (see Egghe and Leydesdorff 2009).

Organization conventions.

> "Each Party shall strive to ensure that such labor principles and the internationally recognized labor rights set forth in Article 16.8 are recognized and protected by its law"(Article 16.1(1) of the 2005 Dominican Republic-Central America Free Trade Agreement; CAFTA-DR)

> "The Parties shall strive to ensure that such labor principles and the internationally recognized labor rights set forth in paragraph 6 are recognized and protected by domestic law" (Article 6(1) of the 2000 United States-Jordan Free Trade Agreement; USA-JOR)

Each clause is broken into sequences of 5-grams to calculate the cosine similarity scores for the clauses. The labor clauses from the two PTAs have a cosine similarity score of 0.65 (see Table 1). Doubling the CAFTA-DR text by repeating it once (i.e. "CAFTA-DR*2") affects its cosine similarity score with the USA-JOR text only by 0.01 rounding error. The scores obtained for CAFTA-DR*2 and CAFTA-DR are highly alike, differing only by 0.02 rounding error.[20] Finally, "both clauses", which refers to the addition of a CAFTA-DR clause and a USA-JOR clause, has a score of 0.89 when compared to CAFTA-DR or to USA-JOR.

	CAFTA-DR	**USA-JOR**	**CAFTA-DR*2**	**Both clauses**
CAFTA-DR	1.00			
USA-JOR	0.65	1.00		
CAFTA-DR*2	0.98	0.64	1.00	
Both clauses	0.89	0.89	0.87	1.00

Table 1: Example of cosine similarity scores (based on 5-grams).

Notes: i. We denote the repetition of CAFTA-DR by CAFTA-DR*2=CAFTA-DR+CAFTA-DR. ii. "Both clauses" refers to CAFTA-DR+USA-JOR.

[20]This difference occurs as there are unique 5-grams in CAFTA-DR*2 that link the end of the first appearance of the clause to the beginning of the repeated clause, which are not present in the text CAFTA-DR. If we were to instead use 1-grams, this score would be 1.00.

3.3 Boilerplate networks

We analyze the extent and patterns of boilerplate in the labor and environmental provisions across PTAs using a statistical network approach (for modeling details see section 4). In our networks, each node (i.e. vertex) represents an NTI, and each tie (i.e. edge) connecting each pair of nodes corresponds to a cosine similarity score (showing the extent of boilerplate within each pair) of 0.50 or greater. While cosine similarity scores range from 0 to 1, with higher values corresponding to greater similarity, for modeling reasons, and for ease of interpretation of the network graphs shown below, the use of a threshold criterion is necessary. As a robustness check (see Appendix C, p. 8), we also apply thresholds based on cosine similarity scores of ≥ 0.25 and ≥ 0.75. While these thresholds are arbitrary, they capture a broad cross-section of boilerplate tie strengths.

Figures 1-3 depict three boilerplate networks: the networks of labor and environmental clauses and, for reference, the network of full PTA texts (i.e., the entire body of the text including the NTI clauses). For better readability, we removed "isolates" (agreements without any boilerplate ties) from the figures and added lines to encircle component groupings, which are non-connected clusters of ties. While Figures 1 and 2 display the boilerplate networks for the relevant NTI, Figure 3 depicts the boilerplate network based on the PTAs' full texts. Overall, the figures illustrate that boilerplate is a common characteristic of both PTA texts taken as a whole as well as of specific peripheral clauses, though to varying degrees.

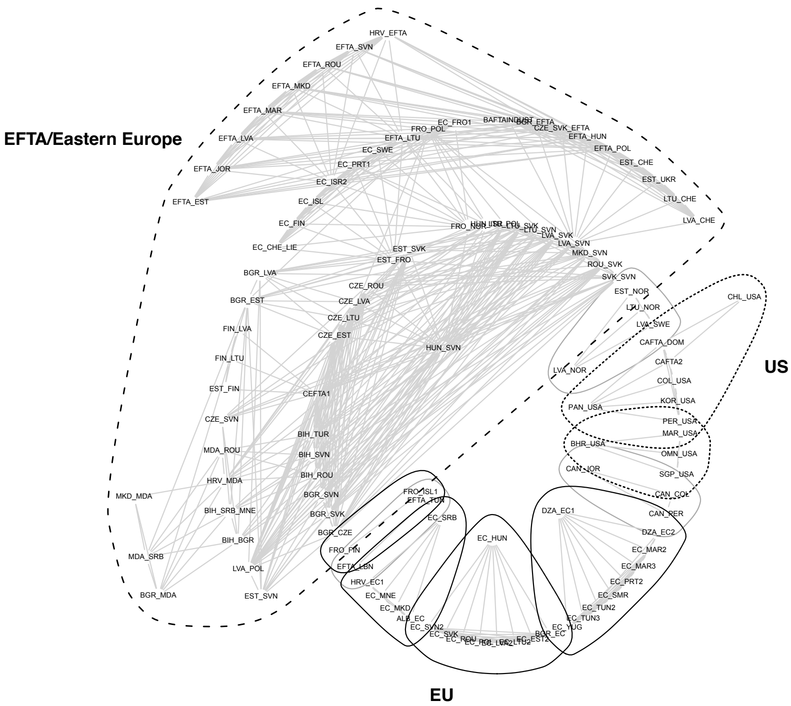

Figure 1: Boilerplate network of PTAs' labor rights clauses.

As Figure 1 illustrates, the labor network includes a variety of actors. The network is divided into ten components, which are the circled non-connected groupings of NTIs. Components outlined in black are discussed below, and the remaining components are outlined in grey. The EFTA/Eastern Europe boilerplate network grouping at the northwest of the diagram is especially large and dense; EFTA is also involved in smaller adjacent grouping. The EU is centrally involved in three groupings at the south of the diagram. US PTAs' labor clauses comprise only two components, one composed of US PTAs with Middle Eastern countries and the other with Latin America. Canada's PTAs and a grouping of Baltic and Nordic PTAs constitute the remaining two components (shown in grey). Taken together, the groupings show that while there is extensive boilerplating on labor issues, it is fairly tightly organized around a few regional groupings.

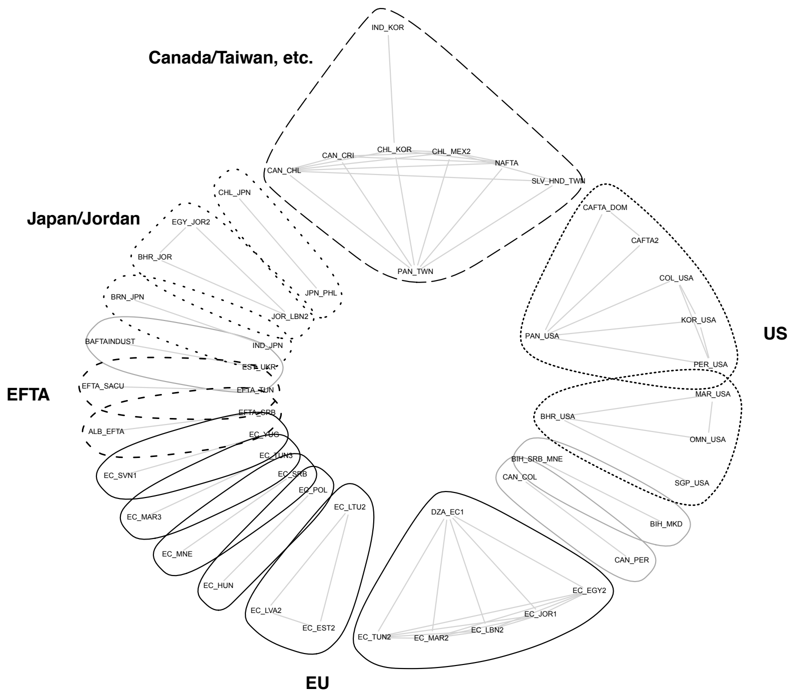

Figure 2: Boilerplate network of PTAs' environmental clauses.

Figure 2 shows that the boilerplate network of PTAs' environmental clauses is less dense. With 17 components, it has a larger number of components than found in the labor network. As in Figure 1, the EU and EFTA, as well as the US and Canada are members of PTAs that feature in the network. However, there are also three components that are exclusively non-European and non-North American. These center on Japanese and Jordanian PTAs. Interestingly, the 1996 Canada-Chile PTA (at west of the top circle in Figure 2) shares boilerplate ties with the 2003 Panama-Taiwan PTA (at south of the top circle) and also the 2007 El Salvador-Honduras-Taiwan agreement (at east of the top circle), showing that boilerplate can occur in the absence of a common member, although that is unusual.

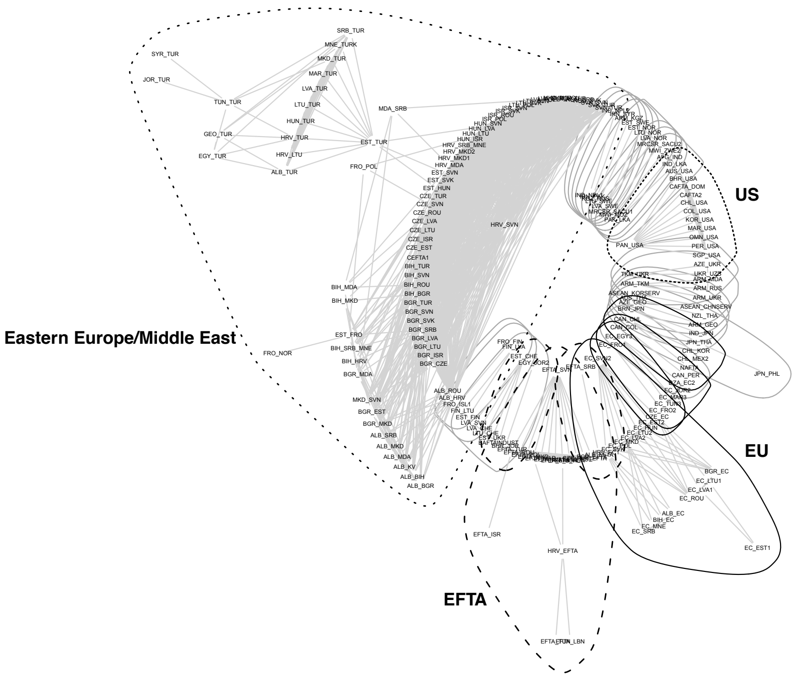

Figure 3: Boilerplate network of PTA full texts.

The full-text network (Figure 3) displays the largest number of components (38), which is unsurprising given the larger number of PTAs. As in Figures 1 and 2, the presence of a common member is a main characteristic of PTAs sharing boilerplate ties. Particularly striking is the dense network of Eastern Europe boilerplating at the northwest of the diagram which is linked through the 1997 Estonia-Turkey PTA to other PTAs between Turkey and both East European and Middle Eastern countries. Ties are also common between the EU and EFTA members, respectively, and between the US and its treaty partners.

Overall, these figures highlight the prominent role of European actors in employing boilerplate across the three networks. However, the cosine similarity score threshold in these figures is set at a minimum of .50 to be consistent with our network analysis and to maximize readability. Accordingly, other states may use boilerplate more or less intensively – we evaluate this possibility using higher and lower thresholds in models shown in Appendix C (p. 8).

Table 2 presents summary statistics for three boilerplate networks. The full-text network, which involves the largest number of PTAs, displays the highest number as well as the highest average number of boilerplate ties (3.68). The lowest average number of boilerplate ties is found in the environment network where each PTA has on average only

0.44 boilerplate ties.

Network	Ties	Mean ties per PTA	Density	PTAs
Labor	470	3.48	0.05	135
Environment	47	0.44	0.01	107
Full text	1281	3.68	0.02	348

Table 2: Summary statistics for boilerplating in PTAs (based on 5-grams).

Notes: Statistics are based on networks with cosine similarity ≥ 0.50.

Density measures the overall intensity of boilerplate in a given network as the number of ties in a given network divided by the possible number of ties. The lowest density scores correspond to the sparsest networks – indicating the least boilerplate, which in this case is found in the boilerplate network of environmental clauses (density=0.01). Conversely, higher density scores indicate more extensive boilerplating, which corresponds to the labor network (density=0.05). Thus boilerplate is much more common in labor provisions than in PTAs taken as a whole and is least common in environmental provisions.

While Table 2 shows that boilerplate is a common feature of PTAs in general, its usage differs across different combinations of states. In the full-text network, the average cosine similarity score for each state's PTA is 0.12. The US' fifteen PTAs have an average cosine similarity score of 0.45 – well above the average – and a total of 117 boilerplate ties (defined by cosine similarity ≥ 0.50). By contrast, Egypt's eleven PTAs have an average cosine similarity score of only 0.02 and none of Egypt's PTAs exceed the 0.50 threshold. The top 20 most similar agreements are presented in Appendix A (p. 1).

4 Estimation technique and model specification

To analyze our NTI boilerplate networks, we use Latent Order Logistic Graph (LOLOG) models (Fellows 2018).[21] LOLOG is a recently developed statistical network technique. Like other similar network approaches (including ERGMs – Exponential Random Graph Models and SAOMs – Stochastic Actor-Oriented Models), LOLOG models treat the interdependence of observations not as a problem but as an opportunity to uncover new information by modeling the structure of the networks. Network models can incorporate system-level as well as PTA-level (dyadic and monadic) covariates to predict overall network characteristics. Their results can be interpreted in manner similar to logistic regression. While LOLOG shares these general advantages of network modeling, it has

[21] These are implemented in R, LOLOG package (Fellows and Handcock 2018).

other distinctive advantages over ERGMs and SAOMs making it particularly suited for analyzing boilerplate.

First, LOLOG models were developed to study growth networks.[22] This makes them especially appropriate for boilerplate where ties between agreements, which represent the presence of shared boilerplate, are created with the signature of a new agreement but do not dissolve or require maintenance post-signing (i.e. as long as legal agreements are not abrogated, the similarity between their texts does not change or diminish over time). Second, LOLOG models incorporate information regarding the order in which nodes enter the network; each potential tie is considered sequentially, beginning with an empty network. Once a PTA enters the network (i.e. has been signed), tie creation between the PTA in question and other PTAs is assumed to operate in a random order. Finally, LOLOG models offer practical benefits: they are generally less prone to degeneracy than ERGMs[23] and are computationally less burdensome than ERGMs and SAOMs. As in logistic regression, the probability of forming a new tie given the network generated up to a particular time point is:

$$\text{logit}(p(y_{s_t} = 1|\eta, y^{t-1}, s_{\leq t})) = \theta \cdot c(y_{s_t} = 1|y^{t-1}, s_{\leq t})$$

where y^{t-1} is the network at time $t-1$, $s_{\leq t}$ is the growth order of the network's ties at time t, θ is a vector of parameters, and $c(y_{s_t} = 1|y^{t-1}, s_{\leq t})$ is a vector that represents the change in network statistics associated with tie creation between time $t-1$ and time t.

Our LOLOG models are fitted using the Monte Carlo Method of Moments (MOM). The resulting untransformed regression coefficient estimates can be interpreted as the change in the likelihood (in log-odds) of tie formation for each unit increase in a covariate. Exponentiated coefficients provide the odds ratio associated with a one-unit increase in a covariate.

4.1 Dependent variable

The networks of NTI boilerplate ties are our dependent variables. Each tie, coded as 1, represents the presence of shared boilerplate across two agreements. The absence of a tie is coded as 0.

[22] SAOMs (see Snijders, Van de Bunt and Steglich 2010) are suited to the analysis of tie creation, maintenance, and dissolution, and ERGMs (see Robins et al. 2007) to model tie creation and dissolution processes.

[23] Degeneracy occurs when a meaningful model cannot be fitted to the data. On degeneracy in ERGMs see Handcock (2003) and Schweinberger (2011). Theoretically, any ERGM distribution can be expressed as a LOLOG and vice versa; in practice, degeneracy or computational constraints may prevent this. When applied to our data, ERGMs displayed strong tendencies towards degeneracy.

4.2 Explanatory variables

Using a series of explanatory variables, we test the extent to which boilerplate in NTI clauses is used for efficiency and/or distribution-related concerns.[24]

As it is empirically impossible to fully disentangle the multiple efficiency-related arguments to account for the declining costs and increasing advantages associated with boilerplate, we include a general efficiency-related term DEGREE POPULARITY.[25] To do so, we use the "geometrically weighted degree" effect with the decay parameter fixed at 1.[26] DEGREE POPULARITY estimates whether particular designs become more popular as they are increasingly used. This degree-related effect captures the change in likelihood of creating an additional tie based on the number of existing boilerplate ties, with marginally decreasing weights for each additional tie given by the decay parameter. Negative coefficient estimates are associated with higher degree NTIs attracting more ties. Thus, NTIs that are popular are more likely to be copied, which has clear efficiency-related benefits such as reducing drafting costs and uncertainty around future legal interpretation.

As a second efficiency-related measure, we include NO COMMON MEMBER, a binary variable coded 1 for a pair of NTIs that does not share any common members. A negative NO COMMON MEMBER coefficient suggests that a boilerplate tie is less likely to be created in NTI pairs that do not share a common member, indicating that states are unlikely to emulate other states' PTA texts without direct contact through a shared member.

To measure the distributive impact of a state's influence on the decision to use boilerplate, we look at combinations of common high- and low-income member states. We first use HIGH-INCOME MEMBER (BOTH), which is coded 1 for NTI dyads where both agreements involve a member classified as a high-income state (World Bank 2018*b*), and 0 otherwise. HIGH-INCOME MEMBER (BOTH) is an influence-related measure as high-income states are more likely to possess the legal and material resources required to draft boilerplate and also the diplomatic ability to press treaty partners to accept their preferred text.[27] As a robustness check, we provide an alternate measure, TOP 20 GDP (BOTH), coded 1 when both agreements include one of the top twenty largest GDPs during any year in the sample and 0 otherwise. Next, we employ COMMON HIGH-INCOME

[24]We deal with missing data using the Amelia method of multiple imputation (Honaker, King and Blackwell 2011).

[25]The term "degree" is a network term that refers to the number of ties a given node has formed.

[26]The geometrically weighted degree statistic counts the nodes in the network with each possible degree value, weights these counts by a decay parameter, and then adds up the weighted counts. The value of the decay parameter is generally chosen for theoretical and/or computational reasons. A value of infinity corresponds to the case where there is no decreasing marginal return as the degree increases. We primarily set the decay value to 1 as it contributes to strong model fit; though as robustness checks, we adjust the decay value to 0.5 and 0 (cf. Hunter 2007).

[27]Wealth is one commonly cited determinant of domestic legal sophistication (Miller 2010); lower income countries are also at a legal disadvantage because of limited economic expertise, bureaucratic resources and their colonial past (Scott and Zahava 2010).

Member, which captures whether an NTI dyad has a common powerful member, thus combining distributive and efficiency-related measures. Common High-Income Member is coded 1 where two PTAs share one or more common high-income members, and 0 otherwise.[28] We also create Common Low-Income Member, which is coded 1 if a pair of PTAs shares a common member not classified as a high-income state (and does not share a common high-income member).[29] The three variables that measure common membership are mutually exclusive and we use Common Low-Income Member as our excluded reference category. Therefore, in models in which we include No Common Member and Common High-Income Member, a positive Common High-Income Member coefficient indicates that NTIs that share a common high-income member are more likely to have boilerplate ties than NTIs with a common member of a lower income level.

Finally, to investigate whether distributive boilerplating preferences are conditioned by substantive issue-area concerns, we account for the most powerful member state's performance regarding the substance of an NTI provision with NTI Performance (Max GDP/cap). In models based on labor rights provisions, NTI Performance (Max GDP/cap) captures the protection of worker rights for the PTA member with the highest (logged) GDP per capita value in the year the agreement was signed.[30] We reverse the coding, with 0 indicating that worker rights are fully protected and 2 that worker rights are severely restricted (Cingranelli, Richards and Clay 2014).[31] A negative coefficient suggests that strong worker rights performance for the most powerful member leads to more boilerplate ties (thereby pointing towards the member's seriousness in promoting protection of labor rights); a positive coefficient, by contrast, suggests that boilerplate is proposed by a powerful member with a weak record of worker rights performance (thus indicating that boilerplate is used as an instrument of window dressing). As a robustness check, we use two alternative measures: NTI Performance (Max GDP) and Physical Integrity Rights Index (Max GDP/cap) (Cingranelli, Richards and Clay 2014) (a reversed scale, ranging from 0 to 8, with lower values corresponding to improved human rights performance).

In models based on the PTA network with environmental provisions, NTI Performance (Max GDP/cap) is coded as a five-year moving average of the annual percent change in greenhouse gas emissions for the member with the highest (logged) GDP per

[28] A state is considered to be high income if it was classified as such at any point during our period of analysis.

[29] There are insufficient observations in each of the income-level categories to facilitate disaggregation into low and middle income categories.

[30] While imperfect, in the area of trade, it is reasonable to assume that the highest income member has an increased ability to influence PTA negotiations to suit its preferences (be it boilerplating or not). Our GDP-based measures come from the Penn World Tables' real GDP output variable (Feenstra, Inklaar and Timmer 2015).

[31] We do so to be consistent with the direction of coding of our environmental performance measure.

capita in the year the PTA was signed.[32] While an imperfect measure of environmental performance, this measure offers the broadest available coverage during our sample years. Again, we can expect two directions of effect regarding a member's environmental performance on boilerplating. As an increase in greenhouse gas emissions indicates poorer environmental performance, a positive coefficient of NTI Performance (Max GDP/cap) in the environmental network implies that a powerful member with a growing emission rate is more likely to use boilerplate as window dressing to avoid higher standards. A negative coefficient, by contrast, indicates that a powerful member with improving environmental performance is more likely to have boilerplate ties (indicating that powerful states use boilerplates as a way of promoting higher environmental standards). As a robustness check, we also use the Environmental Performance Index for the member with the highest (logged) GDP (Esty et al. 2008). However, Environmental Performance Index data is not available for most years in our sample so we use the 2008 index as it is the most recent data available within the timeframe of our analysis. A positive Environmental Performance Index coefficient corresponds to a positive relationship between the environmental performance of the most powerful state within the NTI and boilerplate usage.

4.3 Controls

We also include a number of controls in our analyses. First, we account for the overall level of boilerplate shared between PTAs, by including Full-Text Tie, which is coded 1 if the full texts of a given pair of PTAs have a cosine similarity score ≥ 0.50. Full-Text Tie tests whether NTI boilerplate is related to boilerplate in the rest of a PTA. A positive coefficient indicates that boilerplate choices in the NTI networks are not NTI-specific, but are related to boilerplate choices in the PTA as a whole. Second, the binary variable Same Region is coded 1 if a pair of agreements is from the same region. We expect PTAs within the same region to be more likely to share boilerplate in their NTI clauses due to their shared values and perceived appropriateness as models. Bilateral is a dichotomous variable measuring whether two PTAs are both bilateral to test whether boilerplating depends on the number of states involved. As a robustness check we use Number of Members, a count variable. Finally, Years Apart measures temporal distance in terms of the number of years between each pair of PTAs. We expect PTAs that were signed closer together in time to be more likely to include boilerplate.

[32]NTI Performance (Max GDP/cap) is based on total greenhouse gas emissions in kilotons of CO2 equivalent (World Bank 2018*a*).

	Model 1	Model 2	Model 3	Model 4
Network	*Labor*	*Environment*	*Labor*	*Environment*
Degree Popularity (alpha=1)	-2.039***	-2.068***	-2.071***	-2.178***
	(0.229)	(0.501)	(0.233)	(0.509)
No Common Member	-1.600***	-2.131***	-1.119***	-2.568***
	(0.173)	(0.473)	(0.303)	(0.547)
Common High-Income Member			0.559	-0.659
			(0.314)	(0.545)
High-Income Member (Both)	0.051	0.039	0.066	0.071
	(0.142)	(0.378)	(0.143)	(0.374)
NTI Performance (Max GDP/cap)	0.022	0.067*	0.064	0.055
	(0.089)	(0.026)	(0.095)	(0.028)
Full-Text Tie	2.709***	1.905***	2.698***	2.066***
	(0.145)	(0.423)	(0.150)	(0.453)
Same Region	1.028***	1.039**	1.070***	1.028**
	(0.195)	(0.361)	(0.198)	(0.359)
Bilateral	-0.243*	-0.231	-0.202	-0.209
	(0.119)	(0.303)	(0.126)	(0.299)
Years Apart	-0.220***	-0.144**	-0.218***	-0.146**
	(0.027)	(0.054)	(0.026)	(0.054)
Edges	-0.618	-0.270	-1.231*	0.330
	(0.398)	(1.079)	(0.494)	(1.135)

Table 3: LOLOG models for labor and environmental networks.

Notes: Values are untransformed coefficient estimates with standard errors in parentheses. ***$p \leq .001$, **$p \leq .01$, *$p \leq .05$.

5 Results

In Table 3 and Figure 4, we report the main results of our LOLOG fits for the PTA boilerplate networks with labor (Models 1 and 3) and environmental provisions (Models 2 and 4). We show two sets of models: In Models 1 and 2 we test for the separate effects of efficiency (captured by Degree Popularity and No Common Member) and distributive considerations (High-Income Member (Both)) as well as the issue-area performance for the most powerful PTA member (NTI Performance (Max GDP/cap)). In Models 3 and 4 we also account for the interaction of efficiency and distributive considerations (Common High-Income Member). All estimations are based on cosine similarity scores calculated as 5-grams and with a threshold of ≥ 0.50 (though we vary these assumptions in our robustness checks). Each of our LOLOG models incorporates an edges term, which is a network statistic equal to the number of ties in the network. This term is akin to the intercept in linear regression models (also common to ERGMs).

The results of all our LOLOG models show that efficiency is the prime predictor of boilerplate within the labor and environmental NTI networks. The negative and statistically significant estimates of Degree Popularity indicate that each additional boilerplate tie attached to an NTI further increases the likelihood of future boilerplate ties. This popularity effect suggests that states use boilerplate to exploit the efficiency gains associated with familiar wording. In addition, the negative and statistically significant No Common Member coefficient indicates that agreements that do not share a common member are about 80 percent less likely to share labor boilerplate ties (Model 1) and 88 percent less likely to share environment boilerplate ties (Model 2) than their counterparts. This suggests that emulation of NTIs is unlikely to occur across agreements without direct contact through a shared member.

Turning to distribution-related effects, our analysis reveals that High-Income Member (Both) is not statistically significant in any of the models. Similarly, in Models 3 and 4, Common High-Income Member, which is a combined efficiency and distribution effect, is not statistically significant, indicating that it is a "common" rather than a legally sophisticated "powerful common" member that is conducive to the creation of boilerplate ties. Overall, we find no statistical support for the distribution explanation of boilerplate either alone or in combination with efficiency measures. This may be because these provisions are primarily soft, meaning that they are less effective for purposes of coercion and reflect more their symbolic value for states.

Finally, we find inconsistent support for the combined distribution and substance effect. NTI Performance (Max GDP/cap) is statistically significant only in Model 2, indicating that each unit increase in the most powerful member state's environmental performance score – corresponding to growth in greenhouse gas emissions – increases the odds of an additional boilerplate tie by about 7 percent. This hints that powerful states may use boilerplate as window dressing in areas they prioritize less rather than as a serious way to commit to stricter standards. NTI Performance (Max GDP/cap) is, however, not statistically significant in Models 1, 2 and 4, and is inconsistently significant in other model specifications (see Appendix C, p. 8).

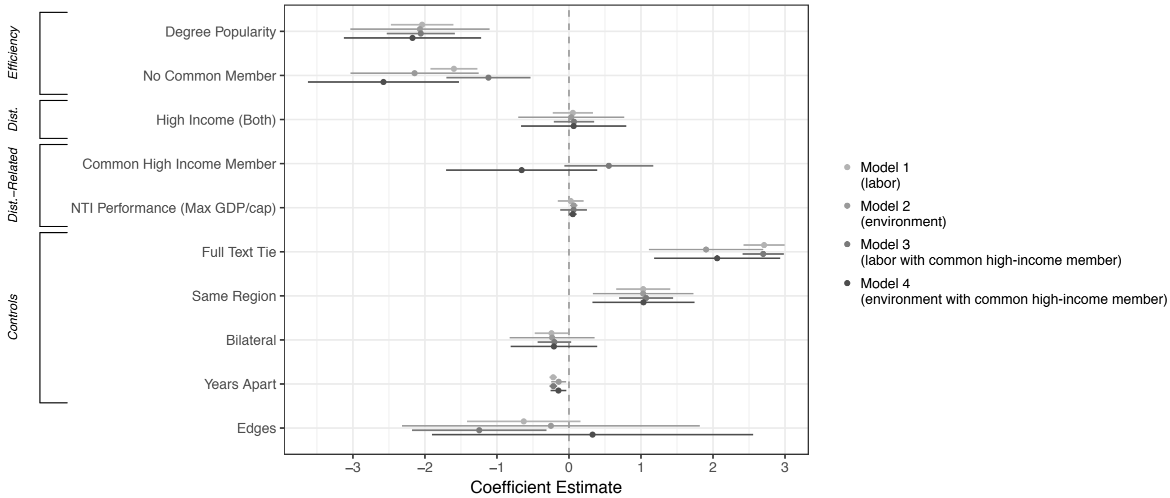

Figure 4: Coefficient plot for parameter estimates from the LOLOG models 1-4.

Our control variables perform largely as expected. Full-Text Tie is positive and statistically significant across all four models. The existence of a full-text tie is associated with a 15-fold increase in the likelihood of creating an NTI tie in the labor network (Model 1), and a 6.7-fold increase in the environmental network (Model 2). These results suggest that NTI boilerplate is not just an NTI-specific feature and is partly explained by PTA-wide considerations. Same Region is positive and statistically significant suggesting that PTAs in the same region are more likely to share boilerplate, which provides an important vehicle for norm diffusion. Bilateral is only statistically significant in Model 1, indicating that bilateral NTI dyads are around 91 percent less likely to share boilerplate than their multilateral counterparts. However, this result is not consistent across the models, limiting its generalizability. Years Apart, in contrast, is as expected negative and statistically significant across all four models, indicating that each additional year that any given two PTAs are apart reduces the likelihood for the two to share boilerplate.

5.1 Goodness of fit

To assess the goodness of fit of our LOLOG models, we report a global network statistic – the degree distribution, which is the distribution of the number of boilerplate ties formed by each PTA.[33] For each model we constructed a plot (see Figure 5, below) to compare the values of this global network statistic in the observed original network to its values in networks that we simulate from our models.

[33] Due to the nature of our data (i.e. a text similarity network), network closure tendencies are strongly determined by the empirical values of the similarity scores rather than by a theory-guided mechanism. For this reason, we are particularly interested in achieving a strong goodness of fit for the degree distribution rather than for the edgewise shared partners distribution - another common measure of goodness of fit. On goodness of fit in LOLOG models see Fellows (2018).

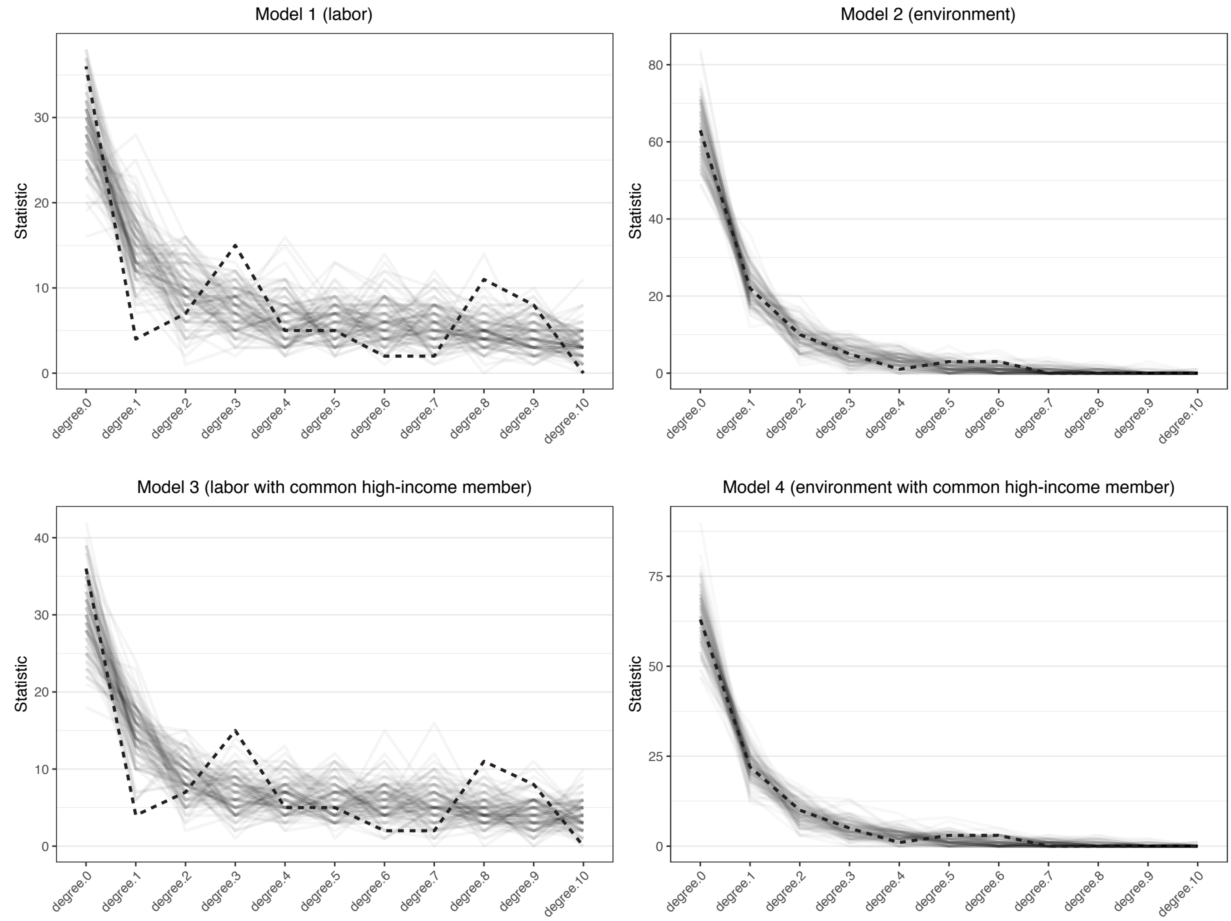

Figure 5: Goodness of fit degree distribution plots for the observed and simulated networks.

Note: Black dotted line represents the observed network, grey lines the simulated networks.

The goodness of fit for all four models is satisfactory. In Models 2 and 4 (environmental networks), there are no clear differences between the observed and simulated networks. In Models 1 and 3 (labor networks), we see some minor statistically significant ($p \leq 0.05$) differences between the simulated and observed values for degrees 1, 3, and 8. This indicates that in these cases we are over- or under-predicting the degrees of nodes and that there is consequently insufficient overlap between the observed and simulated networks. However, as the simulated networks are generally in line with the observed networks in terms of overall degree distribution, as we have no theoretical reason to include additional terms to account for nodes with a degree of 1, 3, or 8 in these particular networks to improve the fit for these degrees, and to facilitate model comparison, we do not introduce further terms into our models. Furthermore, we achieved strong model fits in the labor network in models with the same specification as Model 1 but based on alternate cosine similarity score thresholds of ≥ 0.25 and ≥ 0.75.

5.2 Robustness checks

In Appendix C (p. 8) we show a series of robustness checks confirming that efficiency-related indicators are the most consistent predictors of boilerplate ties.

First, we applied our main model specification from Models 1 and 2 in Table 3 to networks with either lower (≥ 0.25) or higher (≥ 0.75) cosine similarity score thresholds (see Models C1 to C4). The results remain largely consistent with our expectations. NTI Performance is statistically significant in the labor and environmental networks at the lower tie-threshold level (see Models C1 and C2), providing support that relatively powerful states with poorer NTI performance use boilerplate as window dressing. However, this result does not extend to our networks with a higher tie threshold. In addition, at the higher tie threshold (Model C4), Degree Popularity loses its statistical significance.

Second, we re-estimated our main models using extended Jaccard rather than cosine similarity scores (Models C5 and C6). Here we find continued support for the efficiency explanation of boilerplate in the No Common Member coefficients, and inconsistent support for the distribution-related explanation, which varies across the labor and environmental networks. Degree Popularity is not statistically significant in Model C6, likely due to the reduced number of ties present in the extended Jaccard similarity network where repetition within a text affects scores. However, we find that Degree Popularity is negative and statistically significant at the lower threshold of ≥ 0.25 (not shown).

Third, we tested our main models using the cosine similarity threshold of ≥ 0.50 based on 3-grams instead of 5-grams (Models C7 and C8). Our results remain largely robust to our choices of similarity score and n-gram size.

Fourth, we substituted many of the indicators used in the main models with alternates. In Models C9 and C10, we demonstrate that our results are robust to changing the decay parameter of Degree Popularity to 0.5, and in Models C11 and C12 to a value of 0. In Models C13 and C14, we based our distribution and power-related measure on GDP rather than GDP per capita. Top 20 GDP (Both), is positive and statistically significant in the labor network only, showing limited support for the distribution hypothesis. Thus we find no consistent support for the distribution effect on boilerplate.

In Models C15 and C 16, we substituted Bilateral for Number of Members suggesting that the number of members in a PTA is not a statistically significant predictor of boilerplate. Finally, we varied our measures of NTI performance in Models C17 and C18. Neither alternate measure is statistically significant, providing no additional support for the distribution effect on boilerplate.

6 Conclusion

In an era of international treaty proliferation (especially in the area of trade), it is important to understand which states shape the construction of these legal agreements and for what reasons. Although the use of boilerplate is widespread, it has been an overlooked aspect of international treaty making. In domestic contracts, previous studies indicate that boilerplate is employed for multiple reasons such as entrenching power differentials and maximizing efficiency gains. In the international arena, boilerplate used for these reasons may have important consequences. It may privilege the solutions favored by powerful states at the expense of weaker negotiating parties, making gains related to international cooperation increasingly asymmetrical. It may reduce the cost of international contracting for all parties, both powerful and weak. Widely adopted boilerplate may also counteract legal fragmentation, as it reduces potential rule incompatibility.

Combining novel text analysis tools with recent developments in statistical network analysis, we have provided the first insights into the rationales behind and patterns of boilerplate with regard to labor and environmental clauses in PTAs. Our most important finding is that boilerplate in international agreements is primarily determined by efficiency-related considerations rather than by distributive concerns. New labor and environment-related NTI provisions within PTAs incorporate popular and well-understood clauses from previous agreements. Tried and tested solutions are generally favored and are spread directly through shared members.

Contrary to our expectations and evidence from the domestic context, distributive considerations are poor predictors of the use of boilerplate. We find little evidence that more legally sophisticated states are more likely to use boilerplate than other states. What is more, powerful states' decision to use boilerplate text in their NTI provisions is not consistently related to their own performance in that issue area. Thus, while the use of boilerplate in domestic contexts is the result of multiple dynamics, our findings indicate that in international context, boilerplate is primarily used to reap efficiency gains, such as reducing negotiation and drafting costs, and incorporating accumulated knowledge.

Our findings raise several empirical and theoretical questions for future research. First of all, is the irrelevance of distributive considerations in international trade agreements due to the (largely) soft law nature of labor and environmental PTA clauses, and might distributive issues be more important for the use of more precise and binding boilerplate written using hard law language? Next, while we focused here on NTIs in trade agreements, does the prevalence of boilerplate vary within agreements (e.g., provisions regarding trade in goods versus dispute settlement procedures) or depend on their subject matter (e.g., fishery agreements versus investment treaties)? In addition, is boilerplate less common in emerging issues where actors have less experience and are more uncertain about the nature of the underlying problem, and more common in areas of "low politics" where less is at stake? Finally, how is the usage of boilerplate related to states' preferences regarding treaty design features, such as flexibility and scope? Answers to such questions will not only provide important clues into the art of international treaty making but also into international institutional design more broadly.

References

Abbott, Kenneth W. and Duncan Snidal. 2000. "Hard and Soft Law in International Governance." *International Organization* 54(3):421–56.

Acharya, Amitav. 2013. "The R2P and Norm Diffusion: Towards a Framework of Norm Circulation." *Global Responsibility to Protect* 5(4):466–79.

Allee, Todd and Manfred Elsig. 2016. "Are the Contents of International Treaties Copied-and-pasted?" Working Paper. World Trade Institute, University of Bern.

Allee, Todd, Manfred Elsig and Andrew Lugg. 2017. "Is the European Union Trade Deal with Canada New or Recycled? A Text-as-data Approach." *Global Policy* 8(2):246–52.

Alschner, Wolfgang and Dmitriy Skougarevskiy. 2016. "Mapping the Universe of International Investment Agreements." *Journal of International Economic Law* 19(3):561–88.

Alschner, Wolfgang, Julia Seiermann and Dmitriy Skougarevskiy. 2018. "Text of Trade Agreements (ToTA) – A Structured Corpus for the Text as Data Analysis of Preferential Trade Agreements." *Journal of Empirical Legal Studies* 15(3):648–66.

Alter, Karen J. 2012. "The Global Spread of European Style Courts." *West European Politics* 35(1):135–54.

Arnold, Christian and Berthold Rittberger. 2013. "The Legalization of Dispute Resolution in Mercosur." *Journal of Politics in Latin America* 5(3):97–132.

Baccini, Leonardo, Andreas Dür and Yoram Haftel. 2015. Imitation and Innovation in International Governance: The Diffusion of Trade Agreement Design. In *Trade Cooperation: The Purpose, Design and Effects of Preferential Trade Agreements*, ed. Andreas Dr and Manfred Elsig. Cambridge: Cambridge University Press pp. 167–94.

Baker, Mark. 2005. "No Country Left Behind: The Exporting Of U.S. Legal Norms Under The Guise Of Economic Integration." *Emory International Law Review* 19(3):132182.

Bebchuk, Lucian A. and Richard A. Posner. 2005/2006. "One-Sided Contracts in Competitive Consumer Markets." *Michigan Law Review* 104(5):827–36.

Ben-Shahar, Omri. 2005/2006. "Foreword Boilerplate: Foundations of Market Contracts Symposium." *Michigan Law Review* 104(5):821–26.

Benoit, Kenneth. 2017. "quanteda: Quantitative Analysis of Textual Data. R Package Version v0.9.9.65." Accessed September 5, 2018. ⟨http://quanteda.io⟩.

Böhmelt, Tobias and Gabriele Spilker. 2016. "The Interaction of International Institutions from a Social Network Perspective." *International Environmental Agreements: Politics, Law and Economics* 16(1):67–89.

Börzel, Tanja and Thomas Risse. 2012. "From Europeanisation to Diffusion: Introduction." *West European Politics* 35(1):1–19.

Brown, Chester. 2013. *Commentaries on Selected Model Investment Treaties*. Oxford: Oxford University Press.

Burscher, Björn, Daan Odijk, Rens Vliegenthart, Maarten De Rijke and Claes H. De Vreese. 2014. "Teaching the Computer to Code Frames in News: Comparing Two Supervised Machine Learning Approaches to Frame Analysis." *Communication Methods and Measures* 8(3):190–206.

Checkel, Jeffrey T. 2005. "International Institutions and Socialization in Europe: Introduction and Framework." *International Organization* 59(4):801–26.

Cingranelli, David, David Richards and Chad Clay. 2014. "The CIRI Human Rights Dataset." Accessed August 22, 2016. ⟨http://www.humanrightsdata.com⟩.

Clapp, James E., Elizabeth G. Thornburg, Marc Galanter and Fred R. Shapiro. 2011. *Lawtalk: The Unknown Stories Behind Familiar Legal Expressions*. New Haven, CT: Yale University Press.

Damashek, Marc. 1995. "Gauging Similarity with n-Grams: Language-Independent Categorization of Text." *Science* 267(5199):843–8.

Design of Trade Agreements (DESTA) Database. 2018. Accessed December 5, 2018. ⟨https://www.designoftradeagreements.org⟩.

Drezner, Daniel W. 2001. "On the Balance Between International Law and Democratic Sovereignty." *Chicago Journal of International Law* 2(2):321–36.

Drezner, Daniel W. 2007. *All Politics is Global: Explaining International Regulatory Regimes*. Princeton, NJ: Princeton University Press.

Egghe, Leo and Loet Leydesdorff. 2009. "The Relation Between Pearson's Correlation Coefficient r and Salton's Cosine Measure." *Journal of the American Society for Information Science and Technology* 60(5):1027–36.

Elkins, Zachary, Andrew Guzman and Beth A. Simmons. 2006. "Competing for Capital: The Diffusion of Bilateral Investment Treaties." *International Organization* 60(4):811–46.

Esty, Daniel C., Marc A. Levy, Christine Kim, Alex de Sherbinin, Tanja Srebotnjak and Valentina Mara. 2008. "2008 Environmental Performance Index." New Haven: Yale Center for Environmental Law and Policy.

Feenstra, Robert C., Robert Inklaar and Marcel P. Timmer. 2015. "The Next Generation of the Penn World Table." *American Economic Review* 105(10):3150–82.

Fellows, Ian E. 2018. "A New Generative Statistical Model for Graphs: The Latent Order Logistic (LOLOG) Model. arXiv:1804.04583." Accessed September 5, 2018. ⟨https://arxiv.org/abs/1804.04583⟩.

Fellows, Ian E. and Mark Handcock. 2018. "Latent Order Logistic (LOLOG) Graph Models. R Package Version 1.0." Accessed September 5, 2018. ⟨https://github.com/statnet/lolog⟩.

Gilardi, Fabrizio. 2010. "Who Learns From What in Policy Diffusion Processes?" *American Journal of Political Science* 54(3):650–66.

Goldstein, Judith, Miles Kahler, Robert O. Keohane and Anne-Marie Slaughter. 2000. "Introduction: Legalization and World Politics." *International Organization* 54(3):385–99.

Goodman, Ryan and Derek Jinks. 2013. *Socializing States: Promoting Human Rights Through International Law*. Oxford: Oxford University Press.

Grimmer, Justin and Brandon M. Stewart. 2013. "Text as Data: The Promise and Pitfalls of Automatic Content Analysis Methods for Political Texts." *Political Analysis* 21(3):267–97.

Gulati, Mitu and Robert E. Scott. 2012. *The Three and a Half Minute Transaction: Boilerplate and the Limits of Contract Design*. Chicago, IL.: University of Chicago Press.

Hafner-Burton, Emilie M. 2005. "Trading Human Rights: How Preferential Trade Agreements Influence Government Repression." *International Organization* 59(3):593–629.

Handcock, Mark S. 2003. "Assessing Degeneracy in Statistical Models of Social Networks." Working Paper No. 39. Center for Statistics and the Social Sciences, University of Washington.

Hill, Claire A. 2001. "Why Contracts Are Written In Legalese." *Chicago-Kent Law Review* 77(1):59–85.

Hollway, James and Johan Koskinen. 2016. "Multilevel Embeddedness: The Case of the Global Fisheries Governance Complex." *Social Networks* 44(January):281–94.

Honaker, James, Gary King and Matthew Blackwell. 2011. "Amelia II: A Program For Missing Data." *Journal of Statistical Software* 45(7):1–47.

Hooghe, Liesbet and Gary Marks. 2015. "Delegation and Pooling in International Organizations." *Review of International Organizations* 10(3):305–28.

Hopkins, Daniel J. and Gary King. 2010. "A Method of Automated Nonparametric Content Analysis for Social Science." *American Journal of Political Science* 54(1):229–47.

Hunter, David R. 2007. "Curved Exponential Family Models for Social Networks." *Social Networks* 29(2):216–30.

Jetschke, Anja and Philomena Murray. 2012. "Diffusing Regional Integration: The EU and Southeast Asia." *West European Politics* 35(1):174–91.

Jo, Hyeran and Hyun Namgung. 2012. "Dispute Settlement Mechanisms in Preferential Trade Agreements: Democracy, Boilerplates, and the Multilateral Trade Regime." *Journal of Conflict Resolution* 56(6):1041–68.

Johnston, Alastair Iain. 2007. *Social States: China in International Institutions 1980-2000*. Princeton: Princeton University Press.

Kahan, Marcel and Michael Klausner. 1997. "Standardization and Innovation in Corporate Contracting (or "The Economics of Boilerplate")." *Virginia Law Review* 83(4):713–70.

Keck, Margaret E. and Kathryn Sikkink. 2004. *Activists Beyond Borders: Advocacy networks in International Politics*. New York: Cornell University Press.

Kelley, Judith. 2008. "Assessing the Complex Evolution of Norms: The Rise of International Election Monitoring." *International Organization* 62(2):221–55.

Koremenos, Barbara. 2013. "The Continent of International Law." *Journal of Conflict Resolution* 57(4):653–81.

Koremenos, Barbara, Charles Lipson and Duncan Snidal. 2001. "The Rational Design of International Institutions." *International Organization* 55(4):761–99.

Law, Jonathan and Elizabeth Martin. 2014. *A Dictionary of Law*. 7th ed. Oxford: Oxford University Press.

Lenz, Tobias. 2012. "Spurred Emulation: The EU and Regional Integration In Mercosur and SADC." *West European Politics* 35(1):155–73.

Linos, Katerina and Tom Pegram. 2016. "The Language of Compromise in International Agreements." *International Organization* 70(3):587–621.

March, James G. and Johan P. Olsen. 1998. "The Institutional Dynamics of International Political Orders." *International Organization* 52(4):943–69.

Milewicz, Karolina, James Hollway, Claire Peacock and Duncan Snidal. 2018. "Beyond Trade: The Expanding Scope of the Non-trade Agenda In Trade Agreements." *Journal of Conflict Resolution* 62(4):743–73.

Miller, Meredith R. 2010. "Contract Law, Party Sophistication and the New Formalism." *Missouri Law Review* 75(2):493–536.

Morin, Jean-Frédric and Edward Richard Gold. 2014. "An Integrated Model of Legal Transplantation: The Diffusion of Intellectual Property Law in Developing Countries." *International Studies Quarterly* 58(4):781–92.

Morin, Jean-Frédric, Joost Pauwelyn and James Hollway. 2017. "The Trade Regime as a Complex Adaptive System: Exploration and Exploitation of Environmental Norms in Trade Agreements." *Journal of International Economic Law* 20(2):365–90.

Mosley, Paul, Jane Harrigan and John Toye. 1995. *Aid and Power: The World Bank and Policy-Based Lending*. London: Routledge.

Oppenheim, Lassa. 1955. *International Law. A Treatise. Volume 1, Peace*. 2nd ed. London: Longmans, Green.

Partington, Richard. 2019. "Post-Brexit Trade Partners Ask UK to Lower Human Rights Standards." The Guardian. February 13, 2019. ⟨https://www.theguardian.com/politics/2019/feb/13/post-brexit-trade-partners-ask-uk-to-lower-human-rights-standards⟩.

Peacock, Claire. 2018. "Symbolic Regulation: Human Rights Provisions in Preferential Trade Agreements." DPhil Diss. University of Oxford, 2018. ProQuest (13872905).

Perkins, Richard and Eric Neumayer. 2010. "Does the "California Effect" Operate across Borders? Trading- and Investing-Up in Automobile Emission Standards." *Journal of European Public Policy* 19(2):217–37.

Poulsen, Lauge N. Skovgaard. 2014. "Bounded Rationality and the Diffusion of Modern Investment Treaties." *International Studies Quarterly* 58(1):1–14.

Poulsen, Lauge N. Skovgaard. 2015. *Bounded Rationality and Economic Diplomacy: The Politics of Investment Treaties in Developing Countries*. Cambridge: Cambridge University Press.

Robins, Garry, Pip Pattison, Yuval Kalish and Dean Lusher. 2007. "An Introduction to Exponential Random Graph (p*) Models For Social Networks." *Social Networks* 29(2):173–91.

Schimmelfennig, Frank. 2008. "EU Political Accession Conditionality After the 2004 Enlargement: Consistency and Effectiveness." *Journal of European Public Policy* 15(6):918–37.

Schwartz, Alan and Robert E. Scott. 2003. "Contract Theory and the Limits of Contract Law." *Yale Law Journal* 113(3):541–619.

Schweinberger, Michael. 2011. "Instability, Sensitivity, and Degeneracy of Discrete Exponential Families." *Journal of the American Statistical Association* 106(496):1361–70.

Scott, Shirley V. and Orli Zahava. 2010. "Developing countries and International Law." Oxford: Oxford University Press. Accessed December 5, 2018. ⟨http://oxfordre.com/internationalstudies/view/10.1093/acrefore/9780190846626.001.0001/acrefore-9780190846626-e-54/version/0⟩.

Sebastiani, Fabrizio. 2002. "Machine Learning in Automated Text Categorization." *ACM Computing Surveys* 34(1):1–47.

Simmons, Beth A., Frank Dobbin and Geoffrey Garrett. 2006. "Introduction: The International Diffusion of Liberalism." *International Organization* 60(4):781–810.

Simmons, Beth A. and Zachary Elkins. 2004. "The Globalization of Liberalization: Policy Diffusion in the International Political Economy." *American Political Science Review* 98(1):171–89.

Smith, Henry E. 2006. "Modularity In Contracts: Boilerplate and Information Flow." *Michigan Law Review* 104(5):1175–222.

Snijders, Tom A.B., Gerhard G. Van de Bunt and Christian E.G. Steglich. 2010. "Introduction to Stochastic Actor-based Models for Network Dynamics." *Social Networks* 32(1):44–60.

Spilker, Gabriele and Tobias Böhmelt. 2013. "The Impact of Preferential Trade Agreements on Governmental Repression Revisited." *Review of International Organizations* 8(3):343–61.

Tan, Pang-Ning, Michael Steinbach and Vipin Kumar. 2007. *Introduction to Data Mining*. Boston: Pearson Education Inc.

U.S. Congress. 2001/2002. "H.R.3009 – Trade Act of 2002. 107th Cong., 26 July." Accessed May 5, 2017. ⟨https://www.congress.gov/bill/107th-congress/house-bill/3009⟩.

Vandevelde, Kenneth J. 1992. *United States Investment Treaties: Policy and Practice*. Cambridge: Kluwer Law International.

Vogel, David. 1995. *Trading Up: Consumer and Environmental Regulation in a Global Economy*. Cambridge, MA: Harvard University Press.

Vogel, David. 1997. "Trading Up and Governing Across: Transnational Governance and Environmental Protection." *Journal of European Public Policy* 4(4):556–71.

Vreeland, James R. 2003. *The IMF and Economic Development*. New York: Cambridge University Press.

World Bank. 2018*a*. "Total Greenhouse Gas Emissions in kt of CO2 Equivalent." Accessed September 5, 2018. ⟨http://data.worldbank.org/indicator/EN.ATM.GHGT.KT.CE?end=2009&start=1989&view=chart9⟩.

World Bank. 2018*b*. "World Bank Country and Lending Groups." Accessed September 5, 2018. ⟨https://datahelpdesk.worldbank.org/knowledgebase/articles/906519⟩.

Appendix A. Top 20 most similar agreements by NTI area and full text (calculated on 5-grams).

Bilateral agreements are identified by their members' standard three-letter country codes and plurilateral agreements by their abbreviations (e.g. CAFTA stands for the Central American Free Trade Agreement). Signature dates are in parentheses.

1. PTA Labor Provisions

	PTA-Labor 1	PTA-Labor 2	Cosine Similarity
1	BGR-CZE (1995)	BGR-SVK (1995)	1.000
2	BGR-CZE (1995)	BGR-SVN (1996)	1.000
3	BGR-SVK (1995)	BGR-SVN (1996)	1.000
4	CZE-ROU (1994)	ROU-SVK (1994)	1.000
5	LTU-POL (1996)	LTU-SVN (1996)	1.000
6	EFTA-MAR (1997)	EFTA-MKD (2000)	1.000
7	EFTA-MAR (1997)	EFTA-ROU (1992)	1.000
8	EFTA-MAR (1997)	EFTA-SVN (1995)	1.000
9	EFTA-MKD (2000)	EFTA-ROU (1992)	1.000
10	EFTA-MKD (2000)	EFTA-SVN (1995)	1.000
11	EFTA-ROU (1992)	EFTA-SVN (1995)	1.000
12	EST-CHE (1992)	LTU-CHE (1992)	1.000
13	EST-CHE (1992)	LVA-CHE (1992)	1.000
14	LTU-CHE (1992)	LVA-CHE (1992)	1.000
15	BGR-MDA (2004)	MDA-SRB (2003)	1.000
16	EST-NOR (1992)	LTU-NOR (1992)	1.000
17	EST-NOR (1992)	LVA-NOR (1992)	1.000
18	EST-NOR (1992)	LVA-SWE (1992)	1.000
19	LTU-NOR (1992)	LVA-SWE (1992)	1.000
20	LTU-NOR (1992)	LVA-NOR (1992)	1.000

2. PTA Environmental Provisions

	PTA-Environment 1	PTA-Environment 2	Cosine Similarity
1	EC-MAR3 (1996)	EC-TUN3 (1995)	1.000
2	BAFTAINDUST (1993)	EST-UKR (1995)	1.000
3	ALB-EFTA (2009)	EFTA-SRB (2009)	1.000
4	BHR-JOR (2001)	JOR-LBN2 (2002)	1.000
5	BHR-JOR (2001)	EGY-JOR2 (1996)	1.000
6	EGY-JOR2 (1996)	JOR-LBN2 (2002)	1.000
7	CAFTA-DOM (2004)	CAFTA (2004)	0.998
8	COL-USA (2006)	PER-USA (2006)	0.980
9	BHR-USA (2004)	OMN-USA (2006)	0.945
10	CAN-CHL (1996)	NAFTA (1992)	0.860
11	CAN-CHL (1996)	CHL-MEX2 (1998)	0.852

12	EC-MNE (2007)	EC-SRB (2008)	0.784
13	CAN-CHL (1996)	PAN-TWN (2003)	0.762
14	EC-EST2 (1995)	EC-LVA2 (1995)	0.759
15	EC-LTU2 (1995)	EC-LVA2 (1995)	0.755
16	CAFTA2 (2004)	PAN-USA (2007)	0.745
17	CAFTA-DOM (2004)	PAN-USA (2007)	0.744
18	EC-EST2 (1995)	EC-LTU2 (1995)	0.742
19	CAN-CHL (1996)	CAN-CRI (2001)	0.741
20	CHL-MEX2 (1998)	NAFTA (1992)	0.732

3. PTA Full Text

	PTA 1	**PTA 2**	**Cosine Similarity**
1	HRV-MKD (1997)	HRV-MKD2 (2002)	1.000
2	CZE-ISR (1996)	ISR-SVK (1996)	0.991
3	ALB-BIH (2003)	ALB-MDA (2003)	0.991
4	COL-USA (2006)	PER-USA (2006)	0.990
5	CZE-ROU (1994)	ROU-SVK (1994)	0.980
6	EST-FIN (1992)	FIN-LVA (1992)	0.978
7	CAFTA-DOM (2004)	CAFTA (2004)	0.978
8	CZE-TUR (1997)	SVK-TUR (1997)	0.970
9	EFTA-EST (1995)	EFTA-LVA (1995)	0.967
10	CZE-LVA (1996)	LVA-SVK (1996)	0.966
11	BGR-CZE (1995)	BGR-SVK (1995)	0.964
12	CZE-EST (1996)	EST-SVK (1996)	0.958
13	CZE-LTU (1996)	LTU-SVK (1995)	0.956
14	EUU-MNE (2008)	EUU-SRB (2007)	0.944
15	LTU-CHE (1992)	LVA-CHE (1992)	0.944
16	ALB-ROU (2003)	MKD-ROU (2003)	0.934
17	LTU-SVK (1996)	LVA-SVK (1996)	0.934
18	EUU-LTU2 (1995)	EUU-LVA2 (1995)	0.930
19	EST-CHE (1992)	LVA-CHE (1992)	0.926
20	EUU-LTU1 (1994)	EUU-LVA1 (1994)	0.921

Appendix B. List of removed "stop words".

Note that typos in the list are intentional – these remove words that were spelled incorrectly in the underlying text document.

.htm
.html
a
aannex
acp
afghan
afghanistan
africa
african
agadir
agrarian
agreement
aladi
albania
albanian
algeria
algerian
america
american
an
and
andean
andorra
andorran
angola
angolan
annex
antigua
antiguans
anzcerta
appendix
apta
arab
arabia
area
argentina
argentinean
armenia
armenian
article
asean
asia
asia
asian
association
at
australia
australian
austria
austrian
azerbaijan
azerbaijani
b
bafta
bahamas
bahamian
bahrain
bahraini
baltic
bangkok
bangladesh
bangladeshi
barbadian
barbados
barbudans
botswana
bay
belarus
belarusian
belgian
belgium
belize
belizean
benelux
bengal
benin
beninese
bhutan
bhutanese
bimstec
bolivia
bolivian
bosnia
bosnian
botswana
brazil
brazilian
britain
british
brunei
bruneian
bulgaria
bulgarian
burkina
burkinabe
burma
burmese
burundi
burundian
c
cacm
cafta
cam
cambodia
cambodian
cameroon
cameroonian
canada
canadian
cape
caribbean
caricom
cariforum
carifta
cartagena
ceao
cefta
cemac
central
cepgl
chad
chadian
chapter
chile
chilean
china
chinese
cis
cis

city
coast
colombia
colombian
comesa
common
commonwealth
community
comoran
comoros
conference
congo
congolese
convention
cooperation
costa
cotonou
council
countries
croatia
croatian
cuba
cuban
customs
cypriot
cyprus
czech
d
d8
d8pta
danish
democracy
democratic
denmark
development
djibouti
djibouti
dominica
dominican
dutch
e
eac
eaec
east
eastern
ec
eccm
economic
ecota
ecowas
ecsc
ecuador
ecuadorean
eec
efta
egypt
egyptian
eight
eighteen
eighteenth
eighth
eightieth
eighty
el
eleven
eleventh
emirates
emirian
epa
equatorial
eritrea
eritrean
estonia
estonian
ethiopia
ethiopian
eurasian
europe
european
f
federation
fifteen
fifteenth
fifth
fiftieth
fifty
fiji
fijian
filipino
finland
finnish
first
five
five
for
fortieth
forty
forum
four
fourteen
fourteenth
fourth
france
free
french
g
gabon
gabonese
gafta
gambia
gambian
gcc
georgia
georgian
german
germany
ghana
ghanaian
global
great
greater
greece
greek
grenada
grenadian
grenadines
group
gstp
guam
guatemala
guatemalan
guinea
guinea-bissau
guinea-bissauan
guinean
gulf
guuam
guyana
guyanese
h
haiti
haitian
herzegovina
herzegovinian
honduran
honduras
http-eur-lex.europa.eu-en-treaties-dat-11997e-htm-11997e.html
hundred
hundredth
hungarian

hungary
i
kiribati
iceland
icelander
ii
iii
in
independent
india
indian
indonesia
indonesian
initiative
integration
integration
inter-arab
iran
iranian
iraq
iraqi
ireland
irish
islamic
island
islander
islands
islands
israel
israeli
italian
italy
iv
ivorian
ivory
ix
izmir
j
jamaica
jamaican
japan
japanese
jordan
jordanian
k
kazakhstan
kazakhstani
kenya
kenyan
kingdom
kiribati
kittian
kitts
korea
korea
korean
kosovo
kuwait
kuwaiti
kyrgyz
kyrgyzstan
l
lafta
laia
lakes
lanka
lankan
laos
laotian
latin
latin
latvia
latvian
lebanese
lebanon
leone
leonean
lesotho
liberia
liberian
libya
libyan
liechtenstein
er
lithuania
lithuanian
lome
lucia
lucian
luxembourg
luxembourger
m
macedonia
macedonian
madagascar
maghreb
malagasy
malawi
malawian
malaysia
malaysian
maldivan
maldives
mali
malian
malta
maltese
marinese
marino
market
marshall
marshallese
mauritania
mauritanian
mauritian
mauritius
melanesian
mercosur
mexican
mexico
micronesia
micronesian
million
millionth
moldova
moldovan
monacan
monaco
monetary
monetary
mongolia
mongolian
montenegro
moroccan
morocco
mosotho
motswana
mozambican
mozambique
msg
multisectoral
myanmar
n
nafta
namibia
namibian
nations
nauru
nauruan
negotiations
nepal
nepalese
netherlander

netherlands
nevis
nevisian
new
ni-vanuatu
nicaragua
nicaraguan
niger
nigeria
nigerian
nigerien
nine
nineteen
nineteenth
ninetieth
ninety
ninth
north
northern
norway
norwegian
o
oceania
oecs
of
oic
oman
omani
on
one
or
organisation
organization
p
pacific
pakistan
pakistani
palau
palauan
panama
panamanian
papua
papua
paragraph
paraguay
paraguayan
peru
peruvian
philippines
picta
poland
polish
portugal
portuguese
preferences
preferential
pretas
principe
protocol
pta
ptn
q
qatar
qatari
quito
r
republic
rica
rican
romania
romanian
russia
russian
rwanda
rwandan
s
sacu
sadc
safta
saint
saint
salvador
salvadoran
samoa
samoan
san
sao
saudi
saudi
scheme
scottish
second
senegal
senegalese
serbian
serbia
seven
seventeen
seventeenth
seventh
seventieth
seventy
seychelles
seychellois
sierra
sierra
singapore
singaporean
six
sixteen
sixteenth
sixth
sixtieth
sixty
slovakia
slovakian
slovenia
slovenian
solomon
solomon
somali
somalia
south
southeast
southern
spain
spanish
spearhead
sri
sri
st
states
steel
strategic
sudan
sudan
sudanese
suriname
surinamer
swazi
swaziland
sweden
swedish
swiss
switzerland
syria
syrian
system
t
taiwan
taiwanese
tajik
tajikistan

tanzania
tanzanian
tariff
technical
ten
tenth
thai
thailand
the
the
the
the
third
thirteen
thirteenth
thirtieth
thirty
thousand
thousandth
three
three
timor
timorese
to
tobago
tobagonian
togo
togolese
tome
tomean
tonga
tongan
tps
trade
trans
treaty
treaty
trinidad
trinidadian
tripartite
tunisia
tunisian
turkey
turkish
turkmenistan
tuvalu
tuvaluan
twelfth
twelve
twentieth
twenty
two
u
udao
udeac
udeao
uganda
ugandan
ukraine
ukrainian
union
united
uruguay
uruguayan
uzbekistan
uzbekistani
v
vanuatu
vatican
venezuela
venezuelan
verde
verdean
vi
vietnam
vietnamese
vii
viii
vincent
w
welsh
west
www.
x
xi
xii
xiii
xiv
xix
xv
xvi
xvii
xviii
xx
y
yaounde
yemen
yemenite
z
zambia
zambian
zealand
zealander
zimbabwe
zimbabwean
zone

Appendix C. Robustness checks.

In Appendix C, we provide estimation results based on additional model specifications. We discuss the results of these models in the main text of the paper under robustness checks.

In Table C1, we report our main model specification from Table 3 for the two NTI networks (labor and environment) based on cosine similarity scores of ≥0.25 (Models C1 and C2) and ≥0.75 (Models C3 and C4). As with our main models, cosine similarity scores are calculated based on 5-grams. In Models C5 and C6, we calculate our similarity scores using instead extended Jaccard similarity. In Models C7 and C8, the underlying texts are broken into 3-grams rather than 5-grams before being transformed into feature frequency matrices.

Models presented in Table C2 are based on a cosine similarity score of ≥0.50 and calculated on feature-frequency matrices of 5-grams. In Models C9 to C12, the decay parameter associated with DEGREE POPULARITY is changed first to 0 and then to 0.5. In Models C13 and C14, NTI performance is measure for the PTA member with the highest GDP rather than the highest GDP per capita. In Models C15 and C16, we replaced BILATERAL with the NUMBER OF MEMBERS, which gives the total number of members in each PTA. In Models C17 and C18, we use alternate measures of NTI performance – PHYSICAL INTEGRITY RIGHTS (MAX GDP/CAP) in Model C17 and, the ENVIRONMENTAL PERFORMANCE INDEX (MAX GDP/CAP) in Model C18.

	Cosine similarity ≥0.25		*Cosine similarity ≥0.75*		*Extended Jaccard similarity*		*Cosine similarity, 3-grams*	
Model	**C1**	**C2**	**C3**	**C4**	**C5**	**C6**	**C7**	**C8**
Network	*Labor*	*Environment*	*Labor*	*Environment*	*Labor*	*Environment*	*Labor*	*Environment*
DEGREE POPULARITY (ALPHA=1)	-4.652***	-2.789***	-1.606***	-0.910	-1.916***	-0.291	-4.581***	-2.592***
	(1.146)	(0.409)	(0.299)	(1.220)	(0.297)	(0.908)	(0.529)	(0.491)
NO COMMON MEMBER	-2.473**	-2.655***	-2.241***	-3.761***	-1.627***	-3.275**	-2.340***	-2.206***
	(0.785)	(0.372)	(0.245)	(1.078)	(0.218)	(1.117)	(0.406)	(0.442)
HIGH-INCOME MEMBER (BOTH)	0.149	-0.310	0.720**	-0.721	0.881***	-0.029	0.148	-0.178
	(0.159)	(0.296)	(0.237)	(0.566)	(0.226)	(0.501)	(0.148)	(0.362)
NTI PERFORMANCE (MAX GDP/CAP)	0.281*	0.060*	-0.078	0.075	-0.239	0.175***	0.292*	0.065*
	(0.133)	(0.025)	(0.134)	(0.039)	(0.131)	(0.042)	(0.120)	(0.026)
FULL-TEXT TIE	2.990***	1.969***	2.655***	1.113	3.782***	4.188***	3.479***	2.288***
	(0.653)	(0.325)	(0.261)	(0.763)	(0.266)	(0.800)	(0.322)	(0.413)
SAME REGION	1.839**	0.945***	0.664*	1.263	0.885**	2.034**	1.740***	1.418***
	(0.566)	(0.259)	(0.294)	(0.666)	(0.284)	(0.671)	(0.292)	(0.354)
BILATERAL	0.422*	-0.672**	-0.805***	-0.011	-0.679***	-0.269	0.313	-0.498
	(0.166)	(0.231)	(0.166)	(0.450)	(0.159)	(0.418)	(0.167)	(0.284)
YEARS APART	-0.355**	-0.263***	-0.303***	-0.272*	-0.189***	-0.048	-0.259***	-0.198***
	(0.112)	(0.051)	(0.046)	(0.120)	(0.041)	(0.100)	(0.045)	(0.057)
EDGES	0.594	2.214**	0.082	-2.086	-0.716	-5.818**	0.816	0.976
	(0.943)	(0.787)	(0.542)	(2.479)	(0.490)	(2.020)	(0.639)	(0.991)

Table C1: LOLOG models.

Note: Values are untransformed coefficient estimates with standard errors in parentheses.*** $p \leq 0.001$, ** $p \leq 0.01$, * $p \leq 0.05$.

Model	C9	C10	C11	C12	C13	C14	C15	C16	C17	C18
Network	*Labor*	*Environment*	*Labor*	*Environment*	*Labor*	*Environment*	*Labor*	*Environment*	*Labor*	*Environment*
Degree Popularity (alpha=0)	-2.004***	-1.117***								
	(0.254)	(0.296)								
Degree Popularity (alpha=0.5)			-2.022***	-1.559***						
			(0.233)	(0.371)						
Degree Popularity (alpha=1)					-1.962***	-2.005***	-2.125***	-1.991***	-2.081***	-1.896***
					(0.234)	(0.510)	(0.229)	(0.498)	(0.224)	(0.498)
No Common Member	-1.451***	-2.103***	-1.517***	-2.116***	-1.524***	-1.895***	-1.693***	-2.113***	-1.612***	-2.032***
	(0.177)	(0.460)	(0.171)	(0.457)	(0.169)	(0.480)	(0.177)	(0.464)	(0.169)	(0.451)
High-Income Member (Both)	0.067	0.028	0.059	0.031			0.057	-0.117	0.073	-0.147
	(0.148)	(0.366)	(0.148)	(0.365)			(0.147)	(0.371)	(0.144)	(0.359)
NTI Performance (Max GDP)					1.206***	1.075				
					(0.272)	(0.660)				
Top 20 GDP Member (Both)					-0.043	0.085*				
					(0.095)	(0.034)				
NTI Performance (Max GDP/cap)	-0.003	0.064*	0.010	0.065*			0.031	0.063*		
	(0.099)	(0.027)	(0.094)	(0.027)			(0.094)	(0.027)		
Physical Integrity Rights (Max GDP/cap)									0.038	
									(0.039)	
Environmental Performance Index (Max GDP/cap)										-0.020
										(0.023)
Full Text Tie	2.948***	1.936***	2.823***	1.903***	2.422***	1.778***	2.682***	1.849***	2.700***	1.738***
	(0.159)	(0.409)	(0.150)	(0.417)	(0.156)	(0.402)	(0.150)	(0.402)	(0.151)	(0.404)
Same Region	1.122***	0.989**	1.066***	1.023**	0.836***	1.026**	1.013***	0.977**	1.126***	1.020**
	(0.209)	(0.356)	(0.206)	(0.347)	(0.189)	(0.357)	(0.196)	(0.358)	(0.208)	(0.355)
Bilateral	-0.273*	-0.199	-0.259*	-0.218	-0.161	-0.290			-0.279*	-0.147
	(0.124)	(0.304)	(0.125)	(0.299)	(0.114)	(0.292)			(0.121)	(0.311)
Number of Members							0.029	-0.067		
							(0.016)	(0.068)		
Years Apart	-0.200***	-0.136**	-0.209***	-0.141**	-0.220***	-0.155**	-0.220***	-0.143**	-0.219***	-0.147**
	(0.028)	(0.051)	(0.028)	(0.053)	(0.027)	(0.056)	(0.028)	(0.055)	(0.027)	(0.052)
Edges	-1.483***	-2.238**	-1.097**	-1.361	-1.473***	-1.157	-1.031**	-0.252	-0.675	2.819
	(0.375)	(0.765)	(0.379)	(0.860)	(0.427)	(1.208)	(0.396)	(0.986)	(0.380)	(3.787)

Table C2: LOLOG models with alternative and additional parameters.

Note: Values are untransformed coefficient estimates with standard errors in parentheses.$^{***}p\leq0.001$, $^{**}p\leq0.01$, $^{*}p\leq0.05$.