Determining the Authorship of Preferential Trade Agreements: A New Technique Using a Supervised Author Topic Model v2.0

Andrew Lugg
University of Maryland

Neil Lund
University of Maryland

Todd Allee
University of Maryland

Manfred Elsig
World Trade Institute
University of Bern

February 2019

Abstract: Preferential trade agreements (PTAs) are the outcome of a complex negotiation process between multiple parties with divergent preferences. Government officials from the parties to a PTA frequently claim that they played a dominant role in writing the agreement. Yet, trade negotiations take place behind closed doors, which means that determining who actually wrote a given PTA, and thus whose interests it best represents, is difficult. Furthermore, commonly used text analysis techniques used in political science are not well suited for use on multi-authored political documents. Therefore, in this paper we propose a new text-as-data method to assess patterns of authorship in a large collection of PTAs. Our novel Supervised Author Topic Model For Treaty Text (SATT) leverages the previous PTAs signed by countries to infer the preferences governments hold over treaty language. From this we can estimate the proportion of a PTA written by each signatory. We apply this technique to a corpus of 493 PTAs and test propositions about the role of power and capacity in the negotiations of trade agreements. Regression analyses of authorship contributions demonstrate that economically powerful countries often play a dominant role in the treaty-writing process. This result confirms existing theoretical approaches that emphasize the role power as a driver of negotiating success, but, importantly, also give us confidence that our approach can be generalized to other multi-authored political documents.

Introduction

Modern trade agreements are about more than just tariff schedules. Preferential Trade Agreements (PTAs) contain hundreds of provisions on everything from geographical indicators to gambling to environmental protection. Indeed, since tariffs on industrial products are historically low, modern trade agreements cover a variety of technical, non-tariff, and behind-the-border provisions. Important data collection and coding efforts, such as the Design of Trade Agreements (DESTA) project, measure variation in these and other commitments across large bodies of PTAs (Dür, Baccini, and Elsig 2014). This has yielded important insights into the nature of commonly included issue areas (Lechner 2016; Allee and Elsig 2016) and provides details on the design of commitments generally (Baccini, Dür, and Elsig 2015). Nonetheless, a fundamental question remains: which states' interests are represented in a given PTA text? In other words, whose decision was it to include a specific issue and whose preferred language was ultimately selected from among the potential alternatives?

Preferential trade agreements (PTAs) are the result of lengthy negotiations that can take months or even years to complete (Lechner and Wüthrich 2018). The process involves bargaining over complex and contentious issues, many of which have profound economic and political consequences for the stakeholders. At the conclusion of trade talks, participating governments declare victory to their domestic audiences, asserting that their country's preferences are written/are reflected in the final agreement. In reality, however, the final language of a given PTA may be composed of text taken predominantly from one side's previous treaties (e.g. Allee and Lugg 2016) or may represent an effort to push into new issue areas (Allee, Elsig, and Lugg 2017). To further complicate matters, countries do not typically release model PTAs, which means there is rarely a baseline treaty from which to judge whether one side dominated the proceedings.

Despite this complex and unobservable process, determining whose interests are manifest in a given PTA is an important task for researchers. Understanding whose preferences PTAs best reflect can shed light on central questions about the role of power and fairness in international law. Realists contend that powerful states write global rules, which suggests that treaty negotiations should favor the most powerful signatories (Gruber 2000). Yet rational design approaches suggest that the nature of the specific cooperation problem will define variation in treaty content (Koremenos 2016). Further, social-constructivist accounts suggest that treaty language is likely to be bound by what is considered appropriate by the parties (Finnemore and Sikkink 1998). Whatever one's perspective, the challenge for researchers is to develop tools to understand the outcome of the treaty-making process. Doing so will lead to a better understanding why states cooperate and what they get out of it.

In this paper we build on advances in computer science on author topic (AT) models to measure the extent to which negotiated PTAs reflect the preferences of their various governments who negotiated them. We employ a novel variation of Latent Dirichlet Allocation (LDA) that calculates the contributions that individual authors make to multi-authored texts (Pratanwanich and Lio' 2014). Using this approach, we infer the topics and topic content governments most commonly include in their PTAs, and then evaluate empirically how closely negotiated treaty text aligns with a government's preferred text. We apply our technique to a corpus of 493 PTAs. After validating the model, we measure the proportion of a given PTA that is composed of language preferred by each negotiating government.

Next, we motivate several theoretical claims seeking to explain which countries are more (or less) successful at writing PTAs. We theorize that counties with more economic leverage and more negotiating expertise are likely to do better in treaty negotiations. Regression analyses of authorship proportions indicate that the economic leverage, measured by the difference in the size of a state's economy relative to the other signatories, is a key determinant of success. The bigger the economy of a country relative to its partners the better it will do. In some cases powerful countries are able to author more than 95% of given PTA. These findings support a view of the treaty negotiation process that highlights the ability of powerful states to push cooperative outcomes towards their preferences. Importantly, however, the finding indicate that it is not just power in the aggregate, but rather the leverage a has vis-à-vis the parties it chooses to negotiate with. Thus, country's success will depend in large part on the power of the party or parties across the table.

Our approach makes several important contributions. Methodologically, we illustrate a novel text-as-data technique that accounts for the multi-authored nature of treaties. Text-as-data techniques currently in use in political science are designed for use on documents, such as manifestos or political speeches, where a single author is assumed to have written the text. In contrast, our approach allows authors' contributions to vary, which better reflects the underlying process of how treaties are negotiated. Second, our technique complements other text-as-data approaches, which rely on exact matching sequences of text and pairwise comparisons (e.g. Allee and Elsig *forthcoming;* Alschner and Skougarevskiy 2017). By contrast, our topic modeling approach flexibly handles subtle changes in the composition of language, which can help researchers better understand evolution and diffusion in large corpora. Finally, our substantive results show that differences in treaty authorship are determined, in large part, by the leverage of negotiating governments. This confirms previous research demonstrating the ways in which powerful actors shape international law, but also suggests that states may perform better or worse based on whom they choose to negotiate with.

Existing Scholarship

The Political Economy of Trade Agreements

Since 1945 approximately 733 PTAs have been negotiated and signed into law (Dür, Baccini, Elsig 2014). A large body of scholarship has sought to understand why states sign PTAs. Economic approaches predominantly focus on their potential to create beneficial trade flows between the signatories (e.g. Baier and Bergstrand 2004; Bhagwati 1992). In contrast, political science approaches stress the interests and characteristics of a variety of actors. For example, some focus on the role of domestic economic actors, such as industry groups and firms and market structures (Milner 1997; Chase 2003; Grossman and Helpman 1995; Dür 2007; Osgood 2018). Whereas, others look to the characteristics of the signatory states, finding that regime type, veto-players and the incentives of leaders all have an impact on PTA formation (e.g. Maggi and Rodriguez-Clare 2007; Mansfield, Milner, and Rosendorff 2002; Mansfield and Milner 2012; Hollyer and Rosendorff 2012). Finally, other approaches explore the strategic rationale behind PTA formation, including competition for market access, forum-shopping, and alliance politics (Baldwin 1993; Baccini and Dür 2012; Mansfield and Reinhardt 2003; Gowa and Mansfield 1993; Mansfield and Bronson 1997).

More recently, scholars are beginning to scrutinize the specific contents of trade agreements by coding variation in the treaty commitments states write into the agreements. This is motivated by a recognition that states include varied provisions in treaty texts (Koremenos et al 2001; Estevadeordal 2009; Kucik 2012; Dür and Elsig 2015). The DESTA project, for example, has manually coded over 620 agreements on over 100 different features, creating treaty level metrics, such as agreement depth and flexibility, that are potentially associated with a variety of political and economic outcomes of interest (Dür, Baccini, and Elsig 2014). Furthermore, DESTA and associated projects have explored the contents of specific issues areas, such as dispute resolution (Allee and Elsig 2016), environmental provisions (Morin, Dür, and Lechner 2018), and other non-trade issues (Lechner 2016). The importance of these approaches is obvious, as it allows researchers to refine existing theory, ask important new questions, and to carry out more precise empirical tests. Thus, understanding the dimensions across which treaty content varies has become a central research agenda in international relations.

Text-as-data Techniques for Analyzing Treaty Text

Nonetheless, there are still important questions that remain unanswered when the focus is on variation in design at the treaty level. Namely, it is difficult to determine the actor(s) who inserted specific language into the agreement. Cognizant of this limitation several papers have turned to text analytic techniques. One approach has been to use document similarity methods and plagiarism detection programs to search for segments of reused text. The assumption being that this allows the researcher to track the movement of text across treaties, which can reveal information about the preferences of individual actors. These measures have been applied to texts such as Native American Treaties (Spirling 2011), bilateral investment treaties (Alschner and Skougarevskiy 2016), and trade agreements (Allee and Lugg 2016; Allee, Elsig, and Lugg 2017), which have yielded important insights. Allee and Elsig (*forthcoming*), find that large sections of PTAs are taken verbatim from other treaties, which suggests a pervasive "copy-pasting" dynamic in treaty making. Others have found strong regional similarities and evidence of diffusion from landmark treaties, such as NAFTA (Alschner, Seiermann, and Skougarevskiy 2017). These approaches are particularly informative when one has reference texts that serve to anchor comparisons or when expert knowledge can help point to useful comparisons.

By necessity, however, measuring variation in the design of PTAs or performing text analysis using similarity measures occurs at the treaty level, whereas some of the most interesting questions are about the individual actors who are party to a given agreement. Trade agreements are multilateral treaties, negotiated and ultimately signed by anywhere between 2 and 193 signatories. Practically, then, a finalized PTA is the result of a bargaining process between sovereign states where the observed outcome is likely to diverge from what one or all of the parties had in mind at the beginning of negotiations. Armed with methods and data cast at the aggregate treaty level – e.g. after the treaty has been negotiated – researchers have not been in a great position to empirically address which actors were most (or least) successful at shaping the final agreement. Thus, we may be unable to say much about who wrote the language and the extent to which a given treaty matches the preferences of any of the signatories across large bodies of treaty text.

Automated text analysis provides a promising solution to the problem of whose preferences are written into multi-authored political documents, such as PTAs. As computational power has grown so too have the text-as-data techniques available to researchers. As Grimmer and Stewart (2013) note, however, one characteristic of text methods is that they are question and domain sensitive. In essence they require the researcher to use judgment to select the most appropriate text model. Treaty text represents a particularly thorny challenge. The primary issue is that treaties contain multiple authors and are the result of an unobserved bargaining process between the authors. Thus, the proportion written by each author is unlikely to be evenly divided among the authors.

Extant text-as-data techniques used in political science, do not provide an adequate solution to the problem of multi-authored treaty text, because they make an implicit or explicit assumption that a single author composed the text[1]. Even

[1] Programs like WORDSCORES (Laver, Benoit, and Garry 2003) and WORDFISH (Slapin and Proksch 2008) scale documents based on the similarity of their respective word counts. This assumption is perhaps not problematic when the goal is to observe preferences across documents or across time, because authorship contributions are either assumed or not a substantively interesting question. However, we are

approaches that more explicitly problematize authorship are designed to discover an unknown author rather than assess author contributions[2]. Thus we need to develop text-as-data models better suited for multi-authored treaty text. In the next section we discuss such a method.

Gauging Treaty Authorship using Topics Models

We contend that a specific class of authorship attribution models, rooted in topic modeling, offers the most promising solution to understanding how parties fare in treaty negotiations. Probabilistic topic models are a family of machine learning algorithms that enable the discovery of latent themes (topics) in collections of text based on an analysis of only the words. The basic intuition is that documents contain a mixture of different topics. Thus, within a given set of documents each document will be composed of a different proportion of each topic. Further, each topic will be composed of words that are unique to that topic. For example, we may find that a collection of trade agreements tend to cover the topics of goods liberalization, services liberalization, public procurement rules, and intellectual property rights, but any given trade agreement may be composed of different proportions of those topics or may neglect to mention some altogether.

The most widely used topic modeling approach is called Latent Dirchlet

---

specifically interested in the variation between author contributions, which requires a different approach.

[2] For example, authorship can be discovered using dictionary techniques. These rely on creating a word lists that can be used as a fingerprint for a given author. They can then be applied to documents with unknown authorship. This has helped discover, among other things, the authorship of the unsigned *Federalists Papers (Mosteller and Wallace 1964)* and the authorship of *A Cuckoo's Calling* to *Harry Potter author J.K. Rowling.*

Allocation or LDA. Pioneered by Blei, Ng, and Jordan (2003) the LDA model is a parametric Bayesian generative model for documents where each document in a corpus is modeled as a mixture of probabilistic topics, and each word in a document is drawn from a distribution of words associated with the topics (see also Blei 2012).[3] The resulting topics offer a lower dimensional representation of a corpus of texts and can be used for classification, automatic indexing, description, and other applications (see Blei 2012). More technically, given a fixed number of topics (*K)* and a collection of documents (*D)* containing words (*W)*, LDA posits a set of *K* multinomial distributions over *W (*Φ)*,* and a set of *D* multinomial distributions over *K (*θ). Documents are generated by selecting a topic from θ, and then sampling a word from the word distribution for that topic. Thus, each document is treated as having a mixture of topics, and documents are composed by repeatedly sampling from a document's topic mixture (see Blei 2012).

Topic model allows the researcher to extract information about the words used in different topics and the proportion of a given document that is composed of a given topic, which are often informative for researchers. One advantage of topic models is that they can be adapted to incorporate document meta-data, which provides the researcher with tools that can be used draw inferences about the sources of variation in topics and topical content[4]. Perhaps the most well known example in political science is the structural topic model (STM), which was created

[3] As an indication of the popularity of this approach in computer science and natural language processing, the Blei et al., 2003 paper has been cited over 17,000 times on Google scholar.

[4] These are often referred to as correlated topic models.

by Roberts, Stewart, Tingley, and Airoldi (2016) and the expressed agenda model created by Grimmer (2010). These approaches allow the user to estimate topic models that include document specific attributes as covariates, which enables focused comparisons between topics and topical content based on document attributes.

Approaches such as STM allow researchers to address a variety of interesting questions[5]. For example, published applications include looking at differences in the way western media portrays Muslim women (Terman 2017), how presidents talk about economic policy (Dybowski and Adammer 2018), and the difference in how survey respondents answer open ended questions (Roberts et al., 2014). Nevertheless, these approaches still assume that documents are effectively single authored. Thus, STM allows the user to make inferences about differences in topic proportions (topic content) across documents, but estimates this by assuming that different authors wrote each of the documents. While useful for a variety of applications, this is not the same as measuring the proportion of a document composed by an individual author to multi-authored treaty text. Being able to address the later question, we argue, is of particular interest for those who study text where an underlying bargaining dynamic among authors is likely to have produced the observed texts.

Theoretical Expectations

[5] STM allows the researcher to observe: 1) differences in the distribution of topics based on the inclusion of metadata (e.g. whether some documents spend more time on one topic relative to another) and, 2) differences between the distribution of words used in topics based on the metadata (e.g. whether characteristics of the author are associated with different words usage within a given topic).

Empirical Design: A New Method to Assess Treaty Authorship

The Supervised Author Topic Model for Treaty Text (SATT)

In this section we propose a topic model that incorporates information on authorship, but allows the documents to be generated by multiple authors who make varying contributions. Our novel contribution builds on a general class of topic models referred to as author topic (AT) models, and adapts significant features proposed in Prantanwanich and Lio (2014). As in conventional LDA, our model assumes that documents are composed of iterative draws from topics. However, while in LDA, the parameter θ is unique to each document in the corpus, the AT model treats θ as an author-specific variable. These author specific mixture weights represent the probability of seeing a topic given an author, rather than the probability of a topic in each document (Rosen-Zvi et al., 2004). Standard AT models assume that authorship is drawn from a uniform distribution, with each author contributing equal amounts to the topics expressed in the document[6].

For our purpose, however, we add an additional latent variable ψ, which represents the probability that a given author contributed to a document. The subsequent steps are identical to the steps followed in other AT models: a topic is sampled from the author's topic distribution, and a word is sampled from the selected topic. This added condition allows for the possibility that some documents

[6] This modeling strategy is likely appropriate for documents that were produced through cooperation, such as multi-authored scientific papers or popular media, but is problematic if we the texts were produced through a bargaining dynamic.

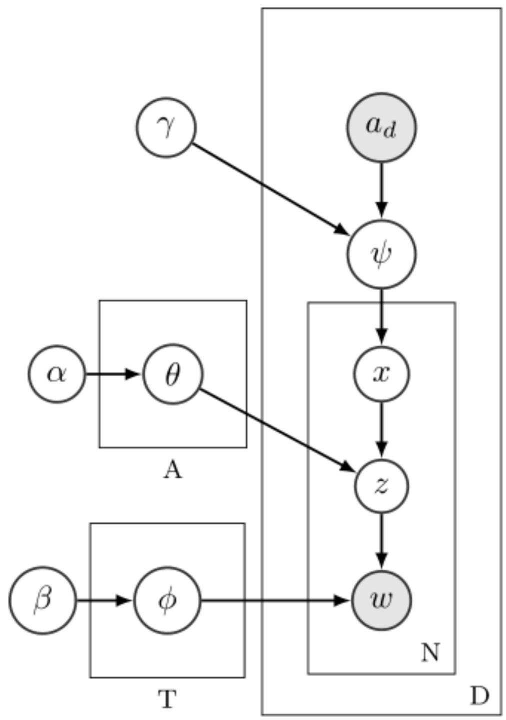
may draw significantly more from the topic preferences of certain authors and significantly less from others. The generative process is outlined below:

1. For each topic $t \in \{1, \dots T\}$
   $a)$ Generate $\phi_t \sim Dir(\beta)$
2. For each author $a \in \{1, \dots A\}$
   $a)$ Choose $\theta_a \sim Dir(\alpha)$
3. For each document $d \in \{1, \dots D\}$
   $a)$ Generate $\psi_d \sim Dirichlet(\gamma a_d)$
4. For each word $i \in \{1, \dots, N_d\}$ in each document $d \in \{1, \dots D\}$
   $a)$ Choose author $x_{d,n} \sim Multinomial(\psi_d)$
   $b)$ Choose topic $z_{d,n} \sim Multinomial(\theta_{x_{d,n}})$
   $c)$ Choose word $w_{d,n} \sim Multinomial(\phi_{z_{d,n}})$

Where ϕ_t is a dirichlet multinomial distribution with a symmetric concentration parameter β. It represents the probability of word w occurring in topic T. θ_a is a dirichlet multinomial distribution with a symmetric concentration parameter α that represents the probability of topic t occurring for author a_d. ψ_d is an asymmetric dirichlet multinomial parameter with a prior parameter $\gamma_{a,d}$. ψ_d represents the probability of an author a occurring in document d. Both $\psi_{a,d}$ and $\gamma_{a,d}$ are set to zero if an author is not a party to a treaty, and are >0 otherwise. Each word in each document is composed by first drawing an author $x_{d,n}$ from ψ_d , then drawing a topic z_{dn} from $\theta_{x,d,n}$, then drawing a word $w_{d,n}$ from $\phi_{zd,n}$.

Figure 1 presents the model in plate notation. Plates represent variables that are repeated. Variables in the outer plate are repeated once for each document, while variables in the inner plate are repeated once for each word in each document. The parameters α, β, and γ are dirichlet priors for θ, Φ, and ψ, respectively. Shaded nodes represent observed data: the author of each document a_d and the words in each document *w,* while clear nodes represent latent variables*.* $x_{d,n}$ is the author assignment for $w_{d,n}$ drawn from ψ_d. $z_{d,n}$ is the topic assignment drawn from the topic distribution $\theta_{x,d,n}$, and $w_{d,n}$ is the observed word drawn from the topic-word distribution $\Phi_{z,d,n}$.

Figure 1. Plate Notation of the SATT model.



Note: Shaded plates are observed data; transparent plates are model parameters and latent variables. Adapted from Prantanwanich and Lio' 2014.

Applying the SATT Model to Treaty Negotiation

In our model the presumed treaty writing process works as follows: each state has a pre-existing set of preferences over topics (θ). For instance, state A might have a strong preference for treaty provisions related to cyber security and agricultural topics, state B might be less interested in cyber security or agriculture, but holds strong preferences for the inclusion of environmental provisions. If state A and state B decide to negotiate a treaty, they will first decide on the scope of negotiations. Next they will negotiate to determine how much of the document they each get to compose (ψ). They will then compose their proportion of the treaty by iteratively drawing words from their topic preference distribution.

State A will frequently draw words from the cyber security and agriculture topics, while state B will frequently draw words from the environmental topic. This generative model is admittedly simplified and abstracted. Still, it conforms to our understanding of the way that negotiations work and is consistent with our conversations with trade negotiators. Furthermore, the conceptualization of the treaty generation process we employ is buoyed considerably by portrayals in the international negotiation literature and (e.g., Lax and Sebenius 1986; Zartman 1988).

The typical PTA, for example, starts with a selection exercise where both parties assess their level of ambition with respect to the topics likely to be included in a potential agreement. We view this as akin to the first stage of our model, where the countries gauge what the author contributions are likely to be. Once formal negotiations begin there are typically successive rounds of negotiations, each of which can last from a few days to several weeks, and which alternate between signatory cities. The average total length of negotiations is about 645 days, but can often be considerably longer (Lechner and Wüthrich 2018).

During the formal rounds, negotiators present draft proposals, which represent that party's preferred language. When there is a lack of agreement the parties typically flag certain passages as needing further deliberation. Between the rounds the negotiators stay in contact by sending each other clarifications, which are then discussed in subsequent round(s). Thus, both sides exchange requests and drafts and the process carries on iteratively until a final agreement is reached. This process matches the way our model is constructed. In an initial phase signatories decide how to divide the topics, then negotiators iteratively negotiate over the amount of topical content (words) that they get to include for a given topic.

Corpus and Model Estimation

We apply the SATT model to a corpus of 493 trade agreements representing 92 different countries. The agreements were collected as part of the DESTA project and represent the most expansive collection of PTA texts currently available (Dür et al., 2014). Annexes and preambles are removed and each text is converted to a machine-readable format. We then used the Stanford Core NLP toolkit to lemmatize and tag the part of speech for each term, and removed all parts of speech other than nouns, verbs, adjectives and adverbs[7]. We then create a term-document matrix that represented each agreement as a series of word-counts.

Our generative process assumes that a set of three latent variables generate agreements: ψ, the portion of the document composed by each author, θ the distribution of each author's preferences over topics, and Φ the distribution of words

[7] Manning, Christopher D., Mihai Surdeanu, John Bauer, Jenny Finkel, Steven J. Bethard, and David McClosky. 2014. The Stanford CoreNLP Natural Language Processing Toolkit In *Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics: System Demonstrations*, pp. 55-60.

used in each topic. Inference proceeds by essentially running this generative process in reverse: we estimate values of ψ, θ, and Φ that are most likely to have generated the observed arrangements of words in documents. As in conventional LDA, these parameters of interest are computationally intractable. Instead we use a Gibbs Sampler to estimate the values of ψ, θ, and Φ.[8] The Gibbs sampler estimates the parameters of interest by continually resampling each author and topic assignment conditional on all other author and topic assignments (see equation 1) until convergence. We randomly assign words to authors and topics, and then, for each word in each document, we calculate the conditional probability that the given word belongs to a topic and author assuming that all other words in the corpus are correctly assigned. We then re-sample author and topic assignment for this word according to this conditional probability, and then move on the next word in the document. After multiple iterations of this process, we eventually converge on a steady state where all of the terms in the corpus have a high probability of being correctly assigned and we can no longer improve our topic or author assignments. We then use this state to infer ψ, θ, and Φ.

[8] The Gibbs sampler was written in C++ and implemented in R using Dirk Eddelbuettel and Romain Francois' Rcpp package. Code is available on request. (check font size for footnotes, 10 or 12

Equation 1

$$P(x_i = a, z_i = t | w_i = v, w_{-i}, x_{-i}, z_{-i}, \alpha, \beta, \gamma, a_i)$$
$$\propto \frac{C^{TV}_{t,v} + \alpha}{\sum_{v'} (C^{TV}_{t,v'} + V\alpha)}$$
$$\times \frac{C^{AT}_{a,t} + \beta}{\sum_{t'} (C^{AT}_{a,t'} + t\beta)}$$
$$\times \frac{C^{DA}_{d,a} + \gamma^{a_{d,a}}}{\sum_{a'} (C^{DA}_{d,a'} + \gamma^{a_{d,a'}})}$$

Model Notation

C^{TV}	T x V matrix of topics by words
C^{AT}	A x T matrix of authors by topics
C^{DA}	D x A matrix of document by author
V, T, A, D	The number of unique words, topics, authors and documents
α, β, γ	Concentration parameters over ϕ, θ, ψ

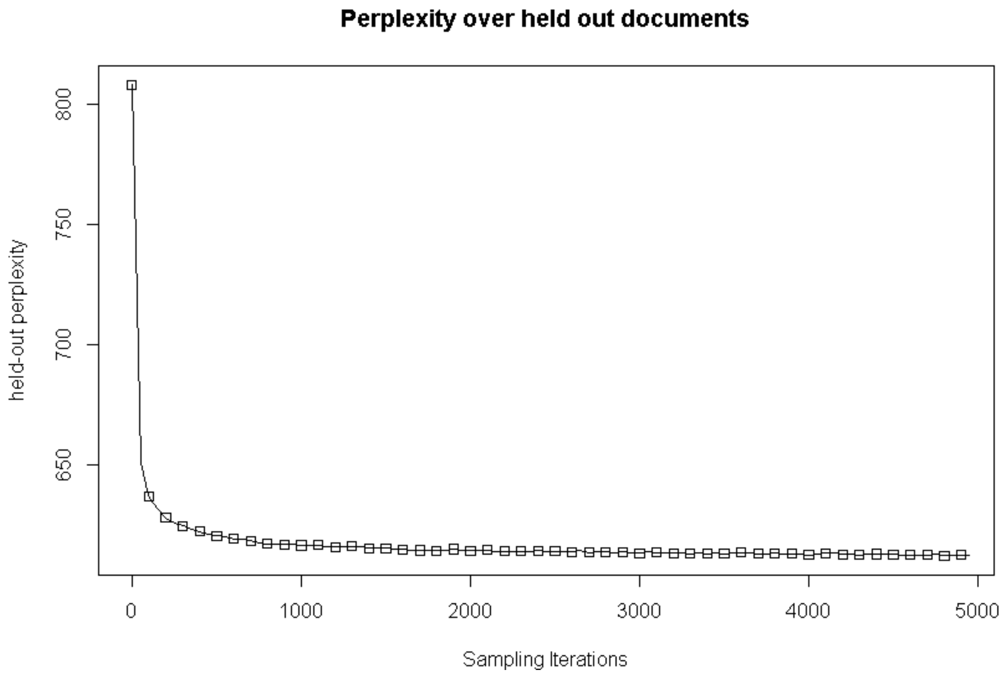
The results below are based on a topic model with K=100 topics, estimated from 5,000 iterations of our Gibbs sampler[9]. We allow the sampler to run for 2,000 iterations in order to allow it to reach convergence (burn-in), and then take samples of ψ, θ, and Φ at every 100th iteration afterward. We use symmetrical priors of

[9] The number of topics was chosen heuristically based on the size of the corpus and vocabulary. In future iterations of the paper we will attempt to find an optimal value of K through a systematic set of model comparisons.

$\alpha=1/K$, $\beta=0.1$, an asymmetric prior $\gamma_{a,d}= 0.1$ for document signatories and $\gamma_{a,d}=0$ otherwise. In order to ensure that the Gibbs sampler reached a steady state, we calculated model perplexity scores at regular sampling intervals for the model. Perplexity is a widely used method for evaluating the performance of text models and is a standard method for assessing topic models (Griffiths and Steyvers 2004). It represents the extent to which a trained model can predict the remaining words in a document after observing some portion of it.[10] Lower values of perplexity indicate that the model is better at predicting held-out words. We held out a random selection of 20% of the words in each document and calculated the model perplexity every 100th iteration of the Gibbs sampler. The results give us confidence that the model reached a steady state: improvement in perplexity level off quickly and show no discernable improvement after 2000 runs.

[10] Perplexity is defined as the (inverse) geometric mean of the summed token likelihoods in test data given the model. A more intuitive explanation is provided by Manning and Shutze (1999; p. 78): "a perplexity of *k* means that you are as surprised on average as you would have been if you had to guess between *k* equiprobable choices at each step." So a perfectly uninformative LDA model would have a perplexity score was equal to the size of the vocabulary of the test set.

Figure 2 Perplexity Scores for SATT Model Runs



Topic Validation

In order to validate the topics we visually inspect the most unique terms for each topic using a frequency-exclusivity (FREX) score adapted from the MALLET toolkit.[11] FREX-scores measure the extent to which a given term is "exclusive" to a topic, and more informative when the same words appear highly probability in several topics simply because they occur with a high frequency in the corpus overall. The FREX scores for six topics are presented in figures 4 – 9 below.

[11] McCallum, Andrew Kachites. "MALLET: A Machine Learning for Language Toolkit." http://mallet.cs.umass.edu. 2002.

Figure 3A. FREX Scores Topic 3

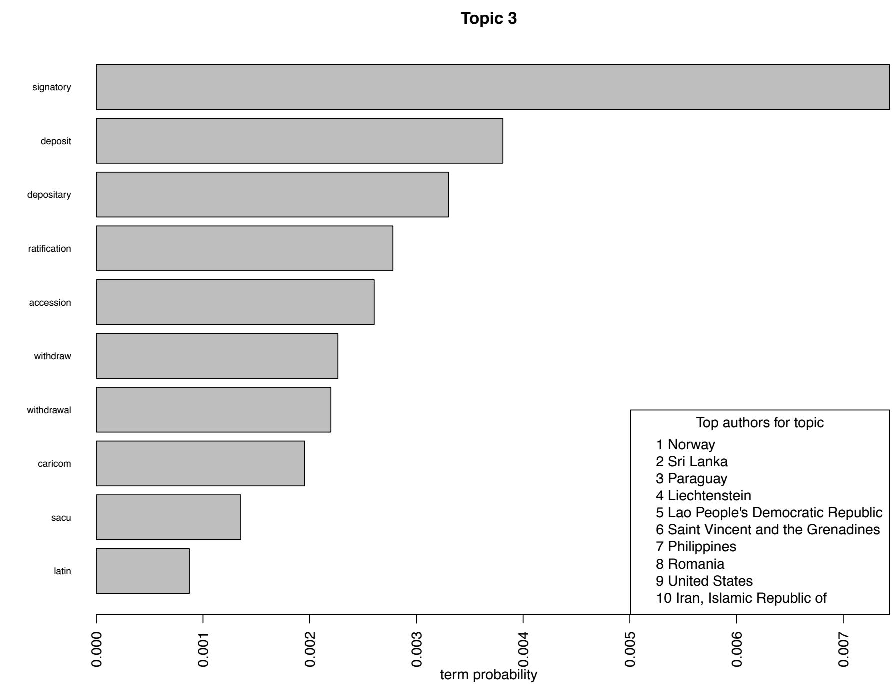


Figure 3B. FREX Scores, Topic 6.

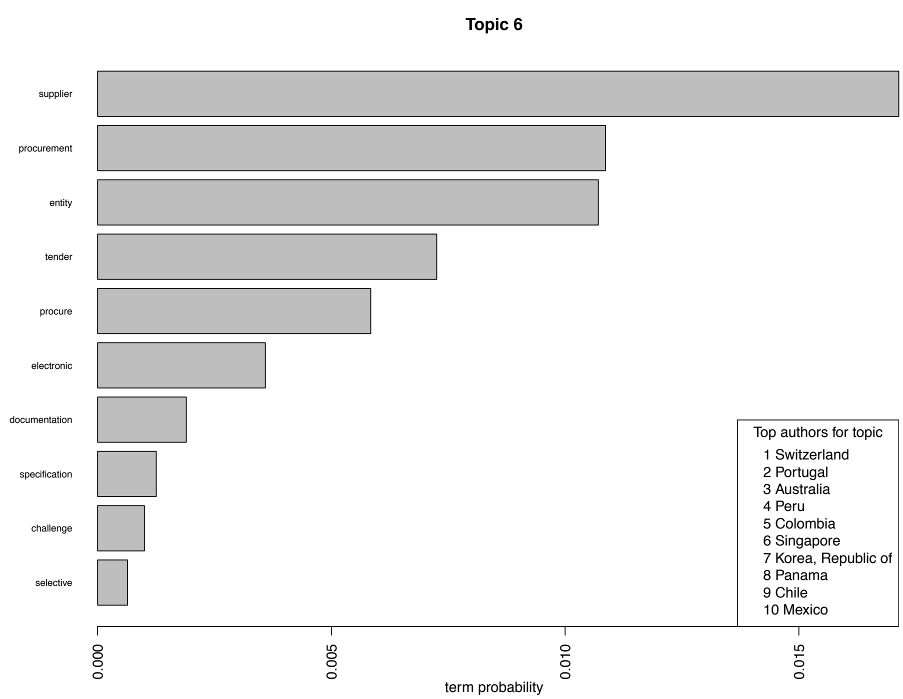


Figure 3C. FREX Scores, Topic 13.

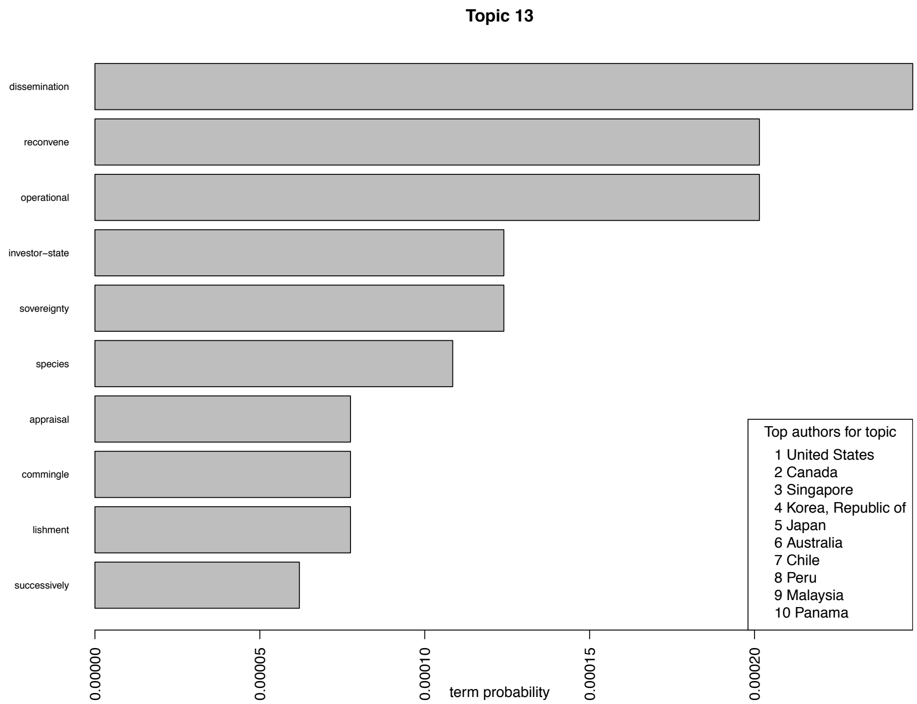


Figure 3D. FREX Scores, Topic 22.

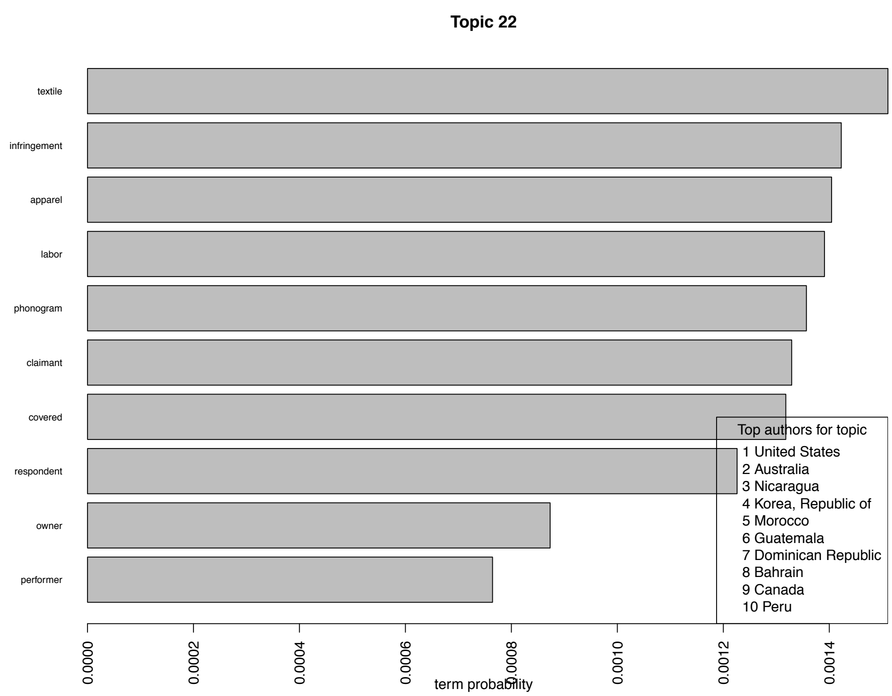


Figure 3E FREX Scores Topic 31.

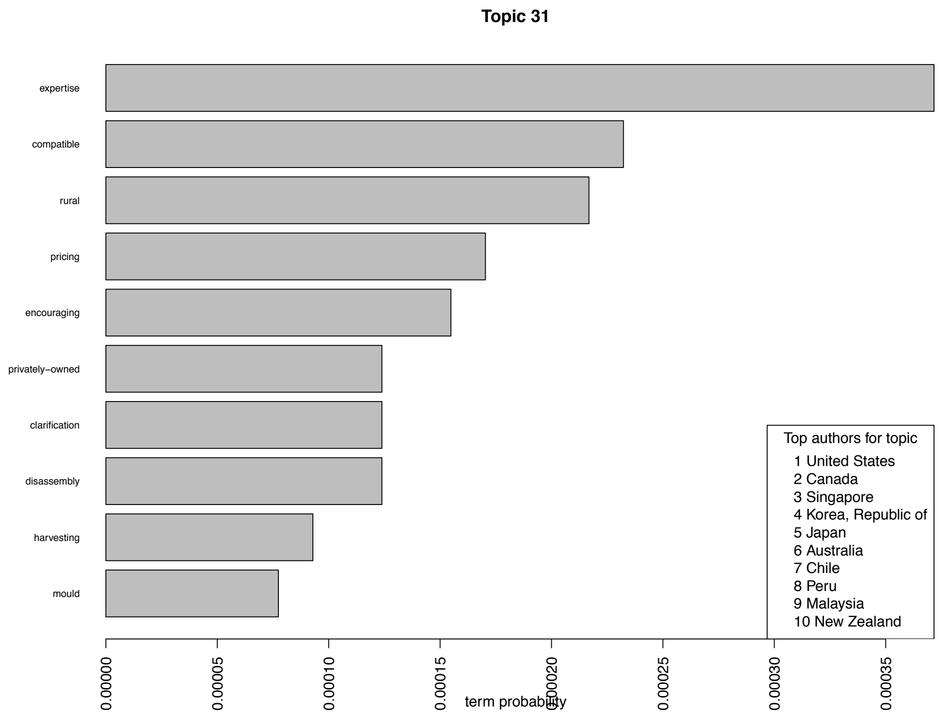
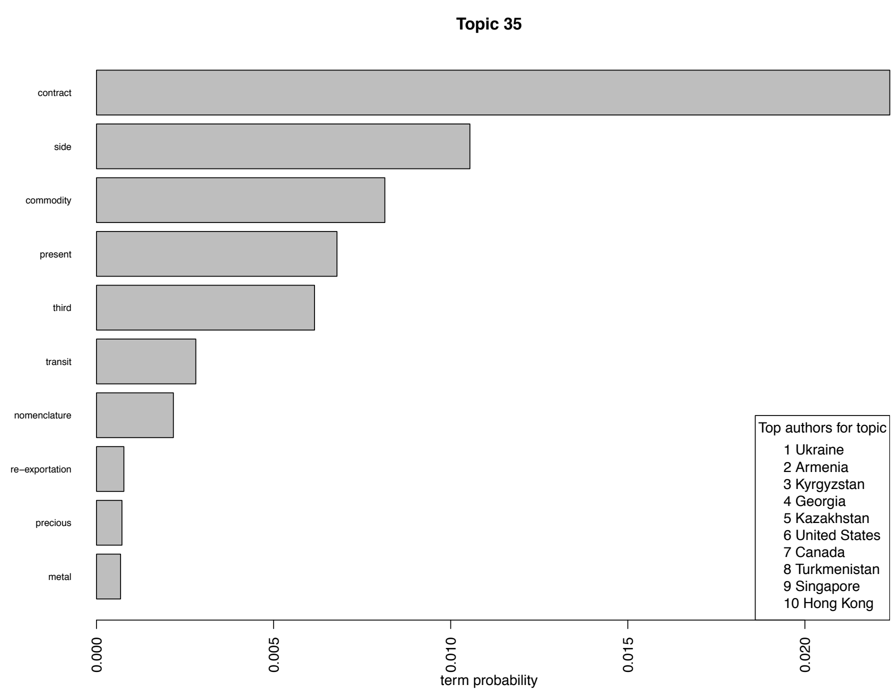


Figure 3F FREX Scores Topic 35.



Topic 6 contains words associated with government procurement, such as "supplier", "procurement", "tender", "procure", and "documentation". Topic 22 has as the top term "textile" and has associated terms like "apparel" and "labor". Topic 35 has words associated with minerals and commodities (e.g. "commodity", "contract", "precious", "metal"), topic 31 contains words associated with agricultural products (e.g. "rural", "harvesting"), topic 13 has terms associated with investor state dispute settlement ("investor state", "sovereignty", "appraisal") and topic 3 has words associated with accession and withdraw (e.g. "signatory", "depositary", "ratification", "succession"). These visual inspections give us confidence that the model is extracting coherent topics. In future versions we will validate the topics statistically using perplexity scores, which is an often- used metric for topic validation.

Empirical Results

Our unique topic modeling approach allows us to test several theoretically informed conjectures about which parties are successful in their treaty negotiations. We have two primary conjectures. Our first is that countries with larger economies, relative to their partners, should be able to embed more of their preferred language into their negotiated treaties. A second conjecture is that countries with higher negotiating capacity in trade should be able to insert more of their preferred language into their negotiated trade agreements.

The estimation of the SATT model on our collection of PTAs produces three primary pieces of information. First, we have a matrix that lists the topics by their respective words counts. Second, we have a matrix of the authors by their respective topics. Finally, we have a matrix of the documents by their respective authors. From these we produce several easily interpretable metrics to gauge the extent to which each signatory to a PTA achieved an authorship role in the final, negotiated text.

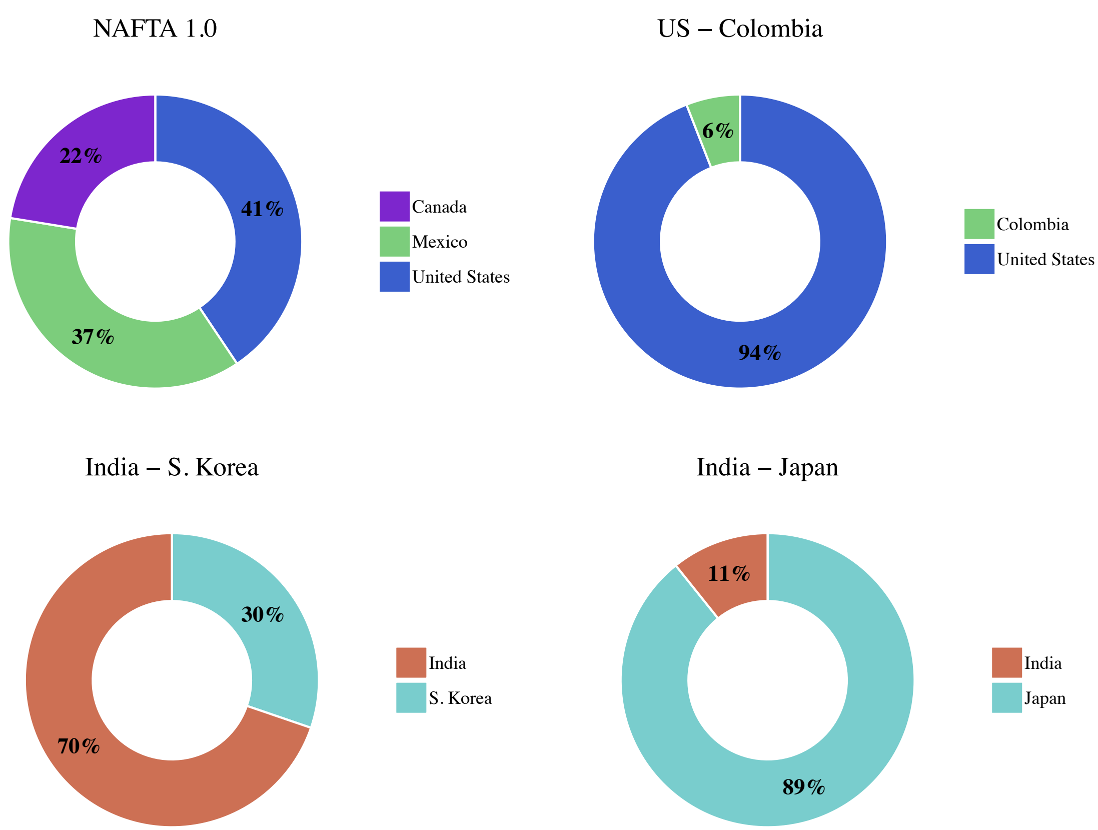
For our first measure we calculate the topic distribution for each document by counting the number of document words assigned to each topic for each author. In essence this represents the number of words that each author gets in the final document from among their preferred topics. We then divide this by the total number of words in each document, which tells us what proportion of the words in a final PTA that come from each author. Naturally, these values range from 0 to 1. Second, we calculate how much this realized authorship proportion differs from what an average signatory should be expected to write based on the number of signatories to each agreement. Thus, a signatory to a 2 party agreement should be expected to split the language 1:1 and a signatory to a 3 party agreement should expect a 1:3. Thus, if a government were to write ¾ (or 75%) of a 2 party agreement, they would receive a score of + .25. These values range from -1 to +1.

As a first step in interpretation we present several salient examples demonstrating variation in authorship proportions across individual PTAs. Fig 4 presents ring charts that summarize the authorship contributions of two US centered and two India centered PTAs.

Focusing on the first PTA in the top row, we see that the author contributions for NAFTA 1.0 were fairly evenly distributed across the three signatory governments. The US composed most of the agreement, approximately 41 percent, but Mexico and

Canada were not far behind, at 37 and 22 percent respectively.

Figure 4 Country Comparisons Across PTAs.

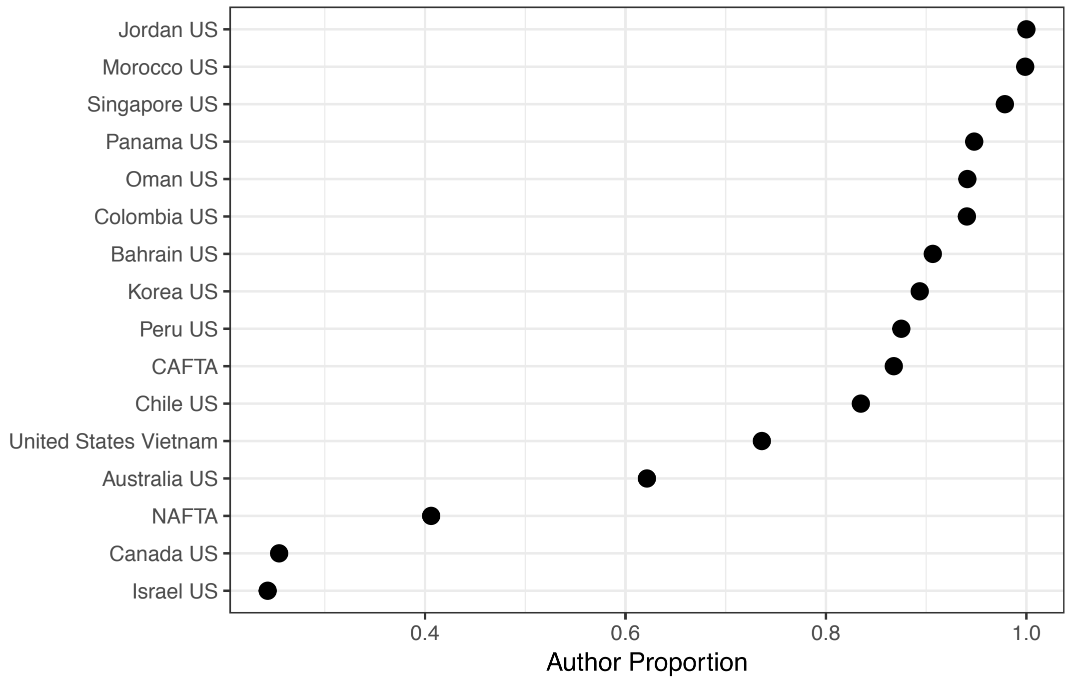


Comparing this to the US Colombia PTA in the second chart on row 1 is illustrative. Here our metrics show that the US composed approximately 94 percent of the agreement text. This difference is likely accounted for by the fact Colombia is much smaller economically and also had not signed many free trade agreements prior to this agreement. Whereas Mexico and Canada combined were much bigger economically and have more experience negotiating trade agreements.

The second row shows two divergent outcomes for India. The first agreement is the India - South Korea PTA from 2009. Here our measure indicates that India was able to author approximately 70 percent of the content, whereas South Korea authored about 30 percent. This makes intuitive sense when one considers that the economy of India was about 60 percent bigger than the Republic of Korea when the agreement was negotiated. However, the second agreement, the India Japan PTA, signed in 2011, shows that India authored only about 11 percent of this agreement compared to 89 percent for Japan. In this case, Japan's economy is over twice as big as India's and Japan, as state with an economy oriented around foreign trade, has skillful and experienced trade negotiators.

We would expect, as the US is one of the world's largest economies that it would tend to play a dominant role writing the majority of its agreements. To empirically explore this the dot plot in figure 5 provides the authorship proportions for all US agreements. As is evident the US tends to perform well when it is paired against much smaller economic partners, including countries such as Jordan, Morocco, and Panama. However, there are several elements that seem to affect US performance. The first is that it performs poorly in early PTAs, such as US Israel, which was the US's first PTA in 1985, and others such as US Canada and NAFTA. Furthermore, we see evidence that larger countries and country groupings, including NAFTA, Australia, and Canada are able to reduce US influence.

Figure 5. US Authorship Performance Across 16 Preferential Trade Agreements



These unique summary measures also allow us to preliminarily test several theoretically informed conjectures about which parties are successful in their treaty negotiations. Our first conjecture is that countries with larger economies should be able to embed more of their preferred language into their negotiated treaties. A second conjecture is that countries with higher negotiating capacity in trade should be able to insert more of their preferred language into their negotiated trade agreements.

Table 1 below estimates a series of OLS regressions at the country-agreement level where the dependent variable is the percentage of the document composed by that author, which can vary between 0 and 100 percent for any given agreement a state is party to. To text our first conjecture – that negotiating leverage drives treaty authorship – we include a measure that calculates the difference between the size of that countries economy and the average size of the other negotiating parties economy (or economy). Thus, positive numbers indicate that a signatory has a larger economy, whereas negative numbers indicate that the signatory has a smaller economy. It is measured in tens of billions of constant 2010 dollars taken from the World Bank's World Development Indicators[12]. We expect that this variable will yield a positive coefficient, indicating that as a country's economy gets bigger than its PTA signatories, it is expected to take a larger authorship role.

Our second conjecture is that countries with greater negotiating capacity than their PTA partners should be able to author more of the final text. To test this conjecture we include a variable that measures the difference in the size of signatories' WTO missions in Geneva. This measure calculates the average WTO mission size of the signatories, and then takes the difference for each partner. Positive numbers indicate a higher negotiating capacity, whereas lower numbers indicate less capacity. This measure is derived from an original annual measure of the raw number of trade representatives each country has at the WTO. We expect the coefficient to positive, which would indicate that countries with higher negotiating capacity should author more of a final PTA text.

We also include several variables meant to control for other potentially confounding factors. First, we include the number of signatories for each agreement. We expect that agreements with a greater number of signatories make it harder for any one author to get what they want in the negotiations. Thus, we expect this coefficient to

[12] We also used GDP in PPP and the GDP from the Penn World Tables. Results were nearly identical.

have a negative sign. Next, we also include country fixed effects, which should help guard against unobserved heterogeneity between the negotiating countries. Finally, we also include a model with agreement fixed effects, which addresses any unobserved heterogeneity across different agreements.

The results from our first set of regressions are presented in table 1 below. Across every model our primary GDP variable yields a positive coefficient that is statistically significant at the 99% level. This indicates that as countries economic size increases relative to their partners, they author a higher percentage of the final document. The level of significance and the coefficient size stay remarkably consistent across the four models we present, which gives us confidence in this result. In addition, included in the appendix, we measure economic size using a variety of other GDP measures, including GDP measured at PPP, and obtain remarkably consistent results. Finally, in addition to be statistically significant, the results are substantively significant as well. Every increase of 10 billion USD in a country's economy yields a half percentage point increase in authorship. This translates to roughly 5 percentage points of increased authorship for every 100 billion dollars. Thus, we see that countries such as the United States are often able to author upwards of 75% or more a given a treaty text.

Models 1 through 4 also include our primary negotiation capacity variable. Although the coefficient is signed correctly, it fails to reach conventional levels of statistical significance across any of our estimated models. One potential issue is that this mission size variable measures a specific kind of trade negotiating capacity that is different from the ability to obtain favorable outcomes in PTA negotiations. Thus, we also included several measures that calculated differences in per capita GDP of the signatories. Unfortunately, these measures were also not statistically significant (see appendix). We also, ran several model specifications where we omit the GDP difference variable to guard against the potential effects of multicollinearity; however, the exclusion of GDP variables does not dramatically affect the results. Thus, we find no initial support for the proposition that negotiation capacity plays a role in determining PTA authorship.

Additionally, models 2 and 3 show that as the number of signatories increases the expected authorship percentage goes down. Most importantly models 3 and 4, which include country and PTA fixed effects, yield similar coefficient estimates for our independent variables, which gives us confidence that the findings are robust to alternative model specifications and potential confounders that we may have failed to include.

Table 1 The Effect of Economic Size and Negotiating Capacity on Authorship Percentages in PTAs.

	Authorship Percentage			
	Model Number			
	(1)	(2)	(3)	(4)
Difference in GDP	0.538^{***}	0.538^{***}	0.465^{***}	0.538^{***}
	(0.105)	(0.096)	(0.139)	(0.115)
Diff in WTO Mission Size	0.205	0.205	0.354	0.205
	(0.337)	(0.310)	(0.415)	(0.370)
Signatories		-2.900^{***}	-3.370^{***}	
		(0.266)	(0.408)	
Constant	36.177^{***}	50.222^{***}	94.284^{***}	50.000^{*}
	(1.475)	(1.873)	(22.190)	(29.269)
Country FE	No	No	Yes	No
PTA FE	No	No	No	Yes
Observations	652	652	652	652
Adjusted R^2	0.061	0.205	0.360	-0.134
Residual Std. Error	37.661 (df = 649)	34.655 (df = 648)	31.103 (df = 527)	41.393 (df = 414)
Note:		*p<0.1; **p<0.05; ***p<0.01		
		Difference in GDP measured in tens of billions of 2010 usd		

Next we run a similar set of OLS regressions where instead of using the raw percentage of a PTA written by each signatory, we create a new dependent variable that calculates how much authorship deviates from what would be expected were each author to contribute equally. This metric determines the deviation, either positive or negative, by dividing 1 by the number of signatories and then taking the difference between this and the observed authorship proportion. Thus, it can theoretically range from -1 to 1, with positive numbers indicating that an author performed better than expected a negative numbers indicating they performed worse. We then multiple this by 100 to aid interpretation. As before we include independent variables to assess the differences economic size and negotiating capacity. However, we do not include a variable for the number of signatories, as this is not longer necessary due to the standardization of the dependent variable. We do, however, include fixed effects at both the country and agreement level.

Table 2 below presents estimates from these models. As before the coefficient measuring the differences in economic size between the signatories is statistically significant at the 99% level across all of the models we estimate. Furthermore, the coefficient estimates yield substantively significant results as well. Based on model 3 we find that a 100 billion dollar increase in GDP relative to a state's negotiating partner(s) yields an increase in the authorship above expectation by almost 6 percentage points.

As before, the primary variable measuring negotiating capacity fails to achieve statistical significance in any of the models we present. We tested different specifications and this time a model with just the WTO difference variable and no other covariates did achieve statistical significance, but it disappears under different specifications. Thus, we again fail to find overwhelming evidence to support the conjecture that negotiation capacity determines authorship patterns in the PTAs we examine.

Table 2 The Effect of Economic Size and Negotiating Capacity on the Authorship of PTAs (Author Deviations from Expected)

	Authorship Above Expected (Percent)		
	Model Number		
	(1)	(2)	(3)
Difference in GDP	0.538***	0.523***	0.538***
	(0.092)	(0.132)	(0.115)
Diff in WTO Mission Size	0.205	0.438	0.205
	(0.296)	(0.399)	(0.370)
Constant	0.032	48.206**	0.000
	(1.295)	(21.378)	(29.269)
Country FE	No	Yes	No
PTA FE	No	No	Yes
Observations	652	652	652
Adjusted R^2	0.079	0.239	-0.444
Residual Std. Error	33.065 (df = 649)	30.060 (df = 528)	41.393 (df = 414)

Note: *p<0.1; **p<0.05; ***p<0.01

Difference in GDP measured in tens of billions of 2010 usd

Conclusions

Our contribution in this paper has been to demonstrate the utility of a novel text analysis technique – the Supervised Author Topic Model for Treaty Text (SATT) – for analyzing political documents that have multiple authors. We applied our technique to a large corpus of PTAs, which have distinctive features that make them difficult to analyze with traditional text-as-data techniques. The model we develop is an improvement on frequently employed text-as-data methods in political science because it explicitly accounts for the underlying generative process of multi-topic and multi-authored documents like PTAs. In particular, our model is an improvement since it allows for variation in the extent to which authors are able, through an underlying process of contestation, to insert their favored language into political documents.

We tested our novel application on a large corpus of preferential trade agreements. After validating our model specification by exploring topics we then generated a dataset that calculated the extent to which a given PTA conformed to that state's preferences, as expressed in the text of the other PTAs they have signed. A descriptive look at the countries that perform well revealed that large trading countries, such as the United States and Japan, do well over the entirety of their trade agreements. Whereas there was also evidence that some small countries may punch above their weight.

We then tested two common conjectures about who does well in international treaty negotiations. Our strongest finding was that states with large economies are able to include more of their preferred text into trade agreements. This is true even after controlling for a variety of other factors such as the depth of the agreement, the number of signatories, and the year of the agreement. We also found evidence that negotiation capacity – as proxied by the size of countries' WTO missions – plays a role. In particular, the results suggest that negotiation capacity may condition the effect of GDP, such that effect of economic size is particularly strong when a state has a high level of legal capacity.

At this point, however, we view the explanatory results as preliminary. There are a variety of other factors that may affect the efficacy of countries in their trade negotiations. Some of these are undoubtedly related to the negotiating context of each individual negotiation, which our models were not able to account for fully. Thus, future iterations will more systematically engage the determinants of negotiating success to include new model specifications. Additionally, there are a variety of steps that can be taken to ensure that the topic model is producing consistent, reliable results. One task is to validate that the findings are robust to a variety of text preprocessing techniques, whereas another is to experiment with different numbers of topics and estimate parameters. An additional step we plan on pursuing is to estimate a model that can account for boilerplate information so as to only focus on substantive topic content.

The novel technique presented here models a common process by which political documents, such as PTAs, are generated. Namely, the texts that political scientists are often interested in are the product of contestation and compromise between several actors that occur through a process of bargaining. This is true not just of international treaties, but also legislative documents and court judgments, to name but a few. As such, we believe that our approach can be used to evaluate a variety of other important research questions across several domains of political science. In particular, our approach could be used to address how authors address topics in a corpus over time and identify documents that are unique (or typical) of that author.

Overall, we view the steps here as an initial effort to apply advanced computational techniques to an area of emerging interest. We view our technique as a complement to manual coding and to text analysis techniques already being employed by scholars studying international treaties in general and PTAs specifically. Our hope is that our results can increase our understanding of the complex process of political contestation that occurs in political documents.

Works Cited

Allee, T., & Elsig, M. (2016). Why do some international institutions contain strong dispute settlement provisions? New evidence from preferential trade agreements. *The Review of International Organizations*, *11*(1), 89-120.

Allee, T., Elsig, M., & Lugg, A. (2017). Is the European Union Trade Deal with Canada New or Recycled? A Text-as-data Approach. Global Policy, 8(2), 246-252.

Allee, T., Elsig, M., & Lugg, A. (2017). The ties between the world trade organization and preferential trade agreements: A textual analysis. Journal of international economic law, 20(2), 333-363.

Allee, T., & Lugg, A. (2016). Who wrote the rules for the Trans-Pacific Partnership?. *Research & Politics*, *3*(3), 2053168016658919.

Allee, T., & Peinhardt, C. (2010). Delegating differences: Bilateral investment treaties and bargaining over dispute resolution provisions. *International Studies Quarterly*, *54*(1), 1-26.

Alschner, W., & Skougarevskiy, D. (2016). Mapping the universe of international investment agreements. *Journal of international economic law*, *19*(3), 561-588.

Alschner, W., & Skougarevskiy, D. (2016). The New Gold Standard? Empirically Situating the Trans-Pacific Partnership in the Investment Treaty Universe. *The Journal of World Investment & Trade*, *17*(3), 339-373.

Alschner, W., Seiermann, J., & Skougarevskiy, D. (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), 648-666.

Baccini, L., Dür, A., & Elsig, M. (2015). The Politics of trade agreement design: revisiting the depth–flexibility Nexus. International Studies Quarterly, 59(4), 765-775.

Baccini, L., & Urpelainen, J. (2014). International institutions and domestic politics: can preferential trading agreements help leaders promote economic reform?. *The Journal of Politics*, *76*(1), 195-214.

Benoit, K., Laver, M., & Mikhaylov, S. (2009). Treating words as data with error: Uncertainty in text statements of policy positions. *American Journal of Political Science*, *53*(2), 495-513.

Blei, D. M. (2012). Probabilistic topic models. *Communications of the ACM*, *55*(4), 77-84.

Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent dirichlet allocation. *Journal of machine Learning research*, *3*(Jan), 993-1022.

Busch, M. L., Reinhardt, E., & Shaffer, G. (2009). Does legal capacity matter? A survey of WTO Members. *World Trade Review*, *8*(04), 559-577.

Büthe, T., & Milner, H. V. (2008). The politics of foreign direct investment into developing countries: increasing FDI through international trade agreements?. American journal of political science, 52(4), 741-762.

Carter, B. E., & Weiner, A. S. (2011). International law. New York: Wolters Kluwer Law & Business.

Chase, K. A. (2009). Trading blocs: States, firms, and regions in the world economy. University of Michigan Press.

Dür, A. (2007). EU trade policy as protection for exporters: The agreements with Mexico and Chile. *JCMS: Journal of Common Market Studies*, *45*(4), 833-855.

Dür, A., & Elsig, M. (Eds.). (2015). Trade Cooperation. Cambridge University Press.

Dür, A., Baccini, L., & Elsig, M. (2014). The design of international trade agreements: Introducing a new dataset. *The Review of International Organizations*, *9*(3), 353-375.

Elkins, Z., Guzman, A. T., & Simmons, B. A. (2006). Competing for capital: The diffusion of bilateral investment treaties, 1960–2000. *International organization*, *60*(04), 811-846.

Fearon, J. D. (1994). Domestic political audiences and the escalation of international disputes. *American Political Science Review*, *88*(03), 577-592.

Fernandez, R., & Portes, J. (1998). Returns to regionalism: An analysis of nontraditional gains from regional trade agreements. The World Bank Economic Review, 12(2), 197-220.

Goldstein, J., Kahler, M., Keohane, R. O., & Slaughter, A. M. (2000). Introduction: Legalization and world politics. *International organization*, *54*(03), 385-399.

Gowa, J., & Mansfield, E. D. (1993). Power politics and international trade. American political science review, 87(2), 408-420.

Griffiths, T. L., & Steyvers, M. (2004). Finding scientific topics. *Proceedings of the National academy of Sciences*, *101*(suppl 1), 5228-5235.

Grimmer, J. (2010). A Bayesian hierarchical topic model for political texts: Measuring expressed agendas in Senate press releases. *Political Analysis*, *18*(1), 1-35.

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

Grossman, G. M., & Helpman, E. (1995). Trade wars and trade talks. Journal of political Economy, 103(4), 675-708.

Gruber, L. (2000). *Ruling the world: Power politics and the rise of supranational institutions*. Princeton University Press.

Hafner-Burton, E. M. (2005). Trading human rights: How preferential trade agreements influence government repression. International Organization, 59(3), 593-629.

Hollyer, J. R., & Rosendorff, B. P. (2012). Leadership survival, regime type, policy uncertainty and PTA accession. International Studies Quarterly, 56(4), 748-764.

Hopkins, D. J., & King, G. (2010). A method of automated nonparametric content analysis for social science. *American Journal of Political Science*, *54*(1), 229-247.

Juola, P. (2008). Authorship attribution. *Foundations and Trends® in Information Retrieval*, *1*(3), 233-334.

Kahler, M. (Ed.). (2015). *Networked politics: agency, power, and governance*. Cornell University Press.

Krasner, S. D. (1999). *Sovereignty: organized hypocrisy*. Princeton University Press.

Laver, M., Benoit, K., & Garry, J. (2003). Extracting policy positions from political texts using words as data. *American Political Science Review*, *97*(02), 311-331.

Lax, D.A. & Sebenius, J.K. (1986). *The manager as negotiator: bargaining for cooperation and competitive gain.* New York: The Free Press.

Koppel, M., Schler, J., & Argamon, S. (2009). Computational methods in authorship attribution. *Journal of the American Society for information Science and Technology*, *60*(1), 9-26.

Koremenos, B. (2013). The continent of international law. *Journal of Conflict Resolution*, *57*(4), 653-681.

Lechner, L. (2016). The domestic battle over the design of non-trade issues in preferential trade agreements. Review of International Political Economy, 23(5), 840-871.

Lechner, L., & Wüthrich, S. (2018). Seal the Deal: Bargaining Positions, Institutional Design, and the Duration of Preferential Trade Negotiations. *International Interactions*, *44*(5), 833-861.

Mansfield, E. D., & Milner, H. V. (2012). *Votes, vetoes, and the political economy of international trade agreements*. Princeton University Press.

Mansfield, E. D., & Reinhardt, E. (2003). Multilateral determinants of regionalism: The effects of GATT/WTO on the formation of preferential trading arrangements. International organization, 57(4), 829-862.

Mansfield, E. D., Milner, H. V., & Rosendorff, B. P. (2002). Why democracies cooperate more: Electoral control and international trade agreements. *International Organization*, *56*(03), 477-513.

Milner, H. V. (1997). *Interests, institutions, and information: Domestic politics and international relations*. Princeton University Press.

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

Morin, J. F., Dür, A., & Lechner, L. (2018). Mapping the Trade and Environment Nexus: Insights from a New Data Set. *Global Environmental Politics*, *18*(1), 122-139.

Mearsheimer, J. J. (2001). *The tragedy of great power politics*. WW Norton & Company.

Pratanwanich, N., & Lio, P. (2014, December). Who wrote this? Textual modeling with authorship attribution in big data. In *Data Mining Workshop (ICDMW), 2014 IEEE International Conference on* (pp. 645-652). IEEE.

Price, R. (1998). Reversing the gun sights: transnational civil society targets land mines. *International organization*, *52*(03), 613-644.

Quinn, K. M., Monroe, B. L., Colaresi, M., Crespin, M. H., & Radev, D. R. (2010). How to Analyze Political Attention with Minimal Assumptions and Costs. *American Journal of Political Science*, *54*(1), 209-228.

Reiter, D., & Stam, A. C. (2002). *Democracies at war*. Princeton University Press.

Roberts, M. E., Stewart, B. M., Tingley, D., Lucas, C., Leder-Luis, J., Gadarian, S. K., ... &

Rand, D. G. (2014). Structural Topic Models for Open-Ended Survey Responses. *American Journal of Political Science*, *58*(4), 1064-1082.

Rosen-Zvi, M., Griffiths, T., Steyvers, M., & Smyth, P. (2004, July). The author-topic model for authors and documents. In *Proceedings of the 20th conference on Uncertainty in artificial intelligence* (pp. 487-494). AUAI Press.

Rosendorff, B. P., & Milner, H. V. (2001). The optimal design of international trade institutions: Uncertainty and escape. International Organization, 55(4), 829-857.

Shaw, M. N. (2008). *International law* (6th ed.). Cambridge England: Cambridge University Press.

Simmons, B. A. (2000). International law and state behavior: Commitment and compliance in international monetary affairs. *American Political Science Review*, *94*(04), 819-835.

Slapin, J. B., & Proksch, S. O. (2008). A scaling model for estimating time-series party positions from texts. *American Journal of Political Science*, *52*(3), 705-722.

Spirling, A. (2012). US treaty making with American Indians: Institutional change and relative power, 1784–1911. *American Journal of Political Science*, *56*(1), 84-97.

Stamatatos, E. (2009). A survey of modern authorship attribution methods. *Journal of the American Society for information Science and Technology*, *60*(3), 538-556

Steinberg, R. H. (2002). In the shadow of law or power? Consensus-based bargaining and outcomes in the GATT/WTO. *International Organization*, *56*(02), 339-374.

Steyvers, M., Smyth, P., Rosen-Zvi, M., & Griffiths, T. (2004, August). Probabilistic author- topic models for information discovery. In *Proceedings of the tenth ACM SIGKDD international conference on Knowledge discovery and data mining* (pp. 306-315). ACM.

Tallberg, J. (2008). Bargaining power in the European Council. *JCMS: journal of common market studies*, *46*(3), 685-708.

Thisted, R., & Efron, B. (1987). Did Shakespeare write a newly-discovered poem?. *Biometrika*, *74*(3), 445-455.

Zartman, I.W. (1988). Common elements in the analysis of the negotiating process. *Negotiation Journal, 4*.