SPRING 2019
NEW YORK UNIVERSITY
SCHOOL OF LAW

"The Making of International Tax Law: Empirical Evidence
from Natural Language Processing"
Omri Marian
University of California at Irvine School of Law

April 2, 2019
Vanderbilt Hall – 208
Time: 4:00 – 5:50 p.m.
Week 10

SCHEDULE FOR 2019 NYU TAX POLICY COLLOQUIUM
(All sessions meet from 4:00-5:50 pm in Vanderbilt 208, NYU Law School)

1. <u>Tuesday, January 22</u> – Stefanie Stantcheva, Harvard Economics Department.

2. <u>Tuesday, January 29</u> – Rebecca Kysar, Fordham Law School.

3. <u>Tuesday, February 5</u> – David Kamin, NYU Law School.

4. <u>Tuesday, February 12</u> – John Roemer, Yale University Economics and Political Science Departments.

5. <u>Tuesday, February 19</u> – Susan Morse, University of Texas at Austin Law School.

6. <u>Tuesday, February 26</u> – Ruud de Mooij, International Monetary Fund.

7. <u>Tuesday, March 5</u> – Richard Reinhold, NYU School of Law.

8. <u>Tuesday, March 12</u> – Tatiana Homonoff, NYU Wagner School.

9. <u>Tuesday, March 26</u> – Jeffery Hoopes, UNC Kenan-Flagler Business School.

10. <u>Tuesday, April 2</u> – Omri Marian, University of California at Irvine School of Law.

11. <u>Tuesday, April 9</u> – Steven Bank, UCLA Law School.

12. <u>Tuesday, April 16</u> – Dayanand Manoli, University of Texas at Austin Department of Economics.

13. <u>Tuesday, April 23</u> – Sara Sternberg Greene, Duke Law School.

14. <u>Tuesday, April 30</u> – Wei Cui, University of British Columbia Law School.

The Making of International Tax Law: Empirical Evidence from Natural Language Processing

Elliott Ash[1] & Omri Marian[2]

We offer the first attempt at empirically testing the level of transnational consensus on the legal language controlling international tax matters. We also investigate the institutional framework of such consensus-building. We build a dataset of 4,052 bilateral income tax treaties, as well as 16 model tax treaties published by the United Nations (UN), Organisation for Economic Co-operation and Development (OECD) and the United States. We use natural language processing to perform pair-wise comparison of all treaties in effect at any given year. We identify clear trends of convergence of legal language in bilateral tax treaties since the 1960s, particularly on the taxation of cross-border business income. To explore the institutional source of such consensus, we compare all treaties in effect at any given year to the model treaties in effect during that year. We also explore whether newly concluded treaties converge towards legal language in newly introduced models. We find the OECD Model Tax Convention (OECD Model) to have a significant influence. In the years following the adoption of a new OECD Model there is a clear trend of convergence in newly adopted bilateral tax treaties towards the language of the new OECD Model. We also find that model treaties published by the UN (UN Model) have little immediate observable effect, though UN treaty policies seem to have a delayed, yet lasting effect. We conclude that such findings support the argument that a trend towards international legal consensus on certain tax matters exists, and that the OECD is the institutional source of the consensus building process.

Key Words: Intergovernmental Organizations, OECD, International Tax, International Law Tax Treaties, Harmonization, Natural Language Processing

JEL Codes: F53, F55, H29, K33, K34, K40

[1] Assistant Professor of Law, Economics, and Data Science, ETH Zurich (ashe@ethz.ch)

[2] Professor of Law and Academic Director of the Graduate Tax Program, University of California, Irvine, School of Law (omarian@law.uci.edu)

I. INTRODUCTION

Whether a binding international legal tax regime exists is an intractable debate defining the academic field of international taxation. The answer to this question has important practical implications. To the extent that there is transnational tax law, nation-states "are not free to adopt any international tax rules they please, but rather operate within the context of the regime" (Avi-Yonah, 2007).

There are, of course, certain formal, binding multilateral tax laws, such as in the European Union (EU), several multilateral tax treaties, and—as of 2016—the Multilateral Convention to Implement Tax Treaty Related Measures to Prevent Base Erosion and Profit Shifting ("MLI").[3] Notwithstanding the existence of these discrete multilateral instruments, the taxation of most cross-border transactions is not subject to any formal transnational legal framework. Rather, most international transactions are taxed in accordance with the laws of the jurisdictions involved, as well about 3,000 bilateral tax treaties. These bilateral tax treaties create the most (some would say only) significant binding transnational legal framework for the taxation of cross-border activities. Some commentators suggest that the network of bilateral tax treaties effectively creates a customary international law of taxation (Avi-Yonah, 2004; Kaufman, 1998). Others reject this notion, arguing that countries are free to adopt whatever tax laws they choose (Rosenbloom, 2000; Roin, 2002).

To the best of our knowledge, no study has empirically investigated the written language of bilateral tax treaties. Such investigation may help to assess whether some level of consensus

---

[3] The MLI is a binding instrument intended to implement certain changes to the tax treaties framework (OECD, 2016). These changes were recommending by the Organization on Economic Cooperation and Development (OECD) as part of its anti Base Erosion and Profits Shifting ("BEPS") project. As of the drafting of this article, 83 jurisdictions have signed on at least certain portions of the MLI.

on international tax matters exists. There has also not been an attempt, to the best of our knowledge, to empirically assess which international institutions has the most influence on the drafting of bilateral tax treaties.

In this article, we make a first attempt to fill this gap in the literature. Using a unique dataset of 4,052 bilateral income tax treaties and 23 model treaties, we seek to empirically answer three questions: First, we try to identify whether consensus on the taxation of cross-border transactions exists as far as controlling legal language is concerned, and whether there are observable changes in the level of such consensus over time. We do so by measuring the level of pair-wise language similarity of each pair of treaties in effect at any given year. We find clear trends towards convergence in legal language in treaties, particularly since the early 1960s — when the OECD first introduced its model tax treaty.

Second, we take advantage of the fact that most bilateral tax treaties are neatly organized into tax topics in order to try and identify whether some areas of international taxation present a higher level of consensus than others. To date, the academic debate on the existence of an international tax regime has run a primarily binary course, with each side of the debate presenting facts supporting its own argument. There was little attempt, however, to empirically identify particular areas of consensus or disagreement. We find that convergence in legal language is most clearly observed in the context of intercompany pricing, taxation of cross-border business income, and in the context of mutual agreement procedures. The lowest levels of convergence are observed in connection with certain definitional issues (such as the taxes and the geographical extent to which treaties apply), on the question on how to relieve double taxation, as well as in the context of assistance in collection of taxes.

Lastly, and perhaps most importantly, we investigate the institutional influence of the drafting of tax treaties. That is, to the extent that there is an international tax consensus embodied in tax treaties, what is its source? The OECD, the UN, and several other international actors issue non-binding model treaties. Commentators point to the fact that actual treaties seem to closely follow the models, particularly the OECD one (Ault, 2009; Ring, 2010). We take advantage of the fact the model treaties are periodically updated, and measure the level of similarity against each model over time. We try to identify whether actual treaties converge in language towards newly published UN or OECD models, or whether such models simply incorporate existing tax treaty practices. We also investigate the model published by the United States, since the United States plays a significant role in affecting international tax policies (Steinmo, 2003; Tanzi, 1995; Kysar, 2016), and we therefore consider the investigation of such a model worthwhile. We find evidence that the OECD Model Tax Convention is most influential. In the years following the adoption of a new OECD Model there is a clear trend of convergence in newly adopted bilateral tax treaties towards the language of the new OECD Model. We find that model treaties published by the UN have little immediate observable effect, but that overall the UN model treaties seem to be similar to existing long-term tax-treaties practices. The models published by the United States tend to respond to existing treaty practices.

While we stop short of concluding that a binding international tax regime exists, we believe our findings lend support to such argument. Trends towards international legal language convergence are clearly observable, at least on certain matters, and IGOs, in particular the OECD, seem to function as institutional sources of consensus-building.

The rest of the article is structured as follows: Section II provides the necessary literary background on two matters. First, it surveys the debate on whether international legal regime

controlling tax matters exists, and the role of bilateral tax treaties in this discourse. Second, we discuss some of the historical background of tax treaties, the rise of model treaties and the institutional framework in which models are concluded. In Section III we discuss the building of the dataset as well as the language processing procedures we use to test legal language convergence. Section IV presents the findings. Section V concludes with a discussion of the implications of our empirical findings.

II. LITERATURE REVIEW

A. Tax Treaties and the "International Tax System"

Some commentators have suggested that there is a coherent international tax regime (Avi-Yonah, 2004; Avi-Yonah, 2007; Brauner, 2003), but opinions differ as to the binding nature of this regime. Pointing to (qualitative) evidence of convergence in domestic tax laws and bilateral tax treaties, some assert that there are identifiable customary norms that provide the basis for a *binding* transnational framework on tax matters (Avi-Yonah, 2004; Kaufman, 1998). Others suggest that the observed convergence in tax standards represents nations' adherence to "soft law" (Brauner, 2016; Christians, 2007; Vega, 2012) or transnational legal orders (Genschel & Rixen, 2015) that are not formally binding, yet create a sense of legal obligation.

National and international governmental bodies also sometime suggest that a coherent international system exists. For example, in response to rampant tax avoidance by multinational corporations and the popular outcry that resulted, the OECD launched the BEPS project in 2013. In part, this project's aim is to revisit "international tax standards" as well as the "international tax framework, which was designed more than a century ago" (Organisation for Economic Cooperation and Development, 2015). Similarly, the U.S. government, in a recent official white

paper, referred to an "international tax system" consisting of certain "international consensus" and "norms" on tax matters (U.S. Department of Treasury, 2016). The white paper was written in response to several decisions by the European Commission. In those decisions, the Commission accused certain EU member states of cutting "sweetheart" tax deals with U.S. corporations. Such deals, according to the commission, amount to illegal state aid under EU law. In a rebuttal attempt, the U.S. Department of Treasury suggested that the European Commission's decisions violate international consensus on certain tax matters.

Other commentators are more skeptical, and suggest that there are no enforceable international tax norms, or in fact, any identifiable "international tax regime" (Rosenbloom, 2000; Roin, 2002). Under such views nations are free to adopt whatever international tax rules they like. According to the skeptical view, international tax law as a field of research deals with the interaction of different *national* tax laws as applied in the cross-border context (Kane, 2004; Rosenbloom, 2000). Specifically, taxpayers engage in cross-border intercompany transactions to reduce their worldwide tax bills. Taxpayers are deliberately taking advantage of differences between *domestic* tax laws across countries in their tax planning ("international tax arbitrage"). Countries, on their part, engage in tax competition through their tax laws in order to attract foreign direct investment. Under the skeptical view, the focus of the academic field of international tax law concerns the policy implications of such behaviors as well as potential responses (if any), whether unilateral or coordinated.

An important factor contributing to the intractability of this debate is the institutional exceptionalism of international taxation. While the bulk of cross-border economic activity is subject to binding multilateral frameworks enforced by centralized intergovernmental organizations (IGOs) such as the World Trade Organization (WTO), there is no formal "world

6

tax organization", nor there has ever been one (Tanzi, 1999). Thus, it is difficult to clearly point to a source of international tax law making, as well as to mechanisms of enforcement (Brauner, 2005).

Instead, taxation of cross-border activity is mostly controlled by a network of thousands of *bilateral* tax treaties. These bilateral legal instruments allocate the taxing rights between the two member states of each treaty. According to recent estimation, there are about 3,000 such treaties in force (Brauner, 2016). In spite of their bilateral nature, tax treaties seem to be remarkably similar to one another. Professor Reuven Avi-Yonah estimated that as much as "75% of the actual words of any given [bilateral tax treaty] are identical with the words of any other [bilateral tax treaty]" (Avi-Yonah, 2009). Since these similar bilateral treaties control the tax treatment of the vast majority of cross-border activity, they arguably constitute an identifiable international tax regime (Brauner, 2016; Avi-Yonah, 2000).

Even if such consensus exists (which many dispute) what might be its source? Is it simply a natural occurrence? Or are there international actors that play a role in the consensus-building process?

The UN and the OECD (as well as other IGOs) publish and update "model" tax treaties from time to time. Such models do not have binding legal force. Rather, they serve as proposed policies by the organizations that issue them. Nonetheless, these model treaties seem have had an identifiable effect on the drafting of bilateral tax treaties (United Nations, 2011). For example, the OECD Model has had a considerable effect on legal standards adopted by OECD members and non-members alike (Ault, 2009; Ring, 2010). Since the 1950s the OECD has emerged as a "solid institutional center" for transnational legal order on tax matters (Genschel & Rixen, 2015).

One commentator suggested that the OECD influence on national tax policies is so profound that the OECD functions as a de-facto "world tax organization" (Cockfield, 2006).

In spite of the significance of these issues, no attempt has been made to date — to the best of our knowledge — to provide systematic empirical evidence of the level of consensus in bilateral tax treaties' language over time, or of the effect of model treaties. While scholars have considered international tax practices from comparative perspectives, past studies have mainly been limited to measurable fiscal outcomes such as tax rates (Slemrod, 2004), revenue compositions (OECD, 2017), the distribution of tax burdens (Swank & Steinmo, 2002), and qualitative assessments of states' tax policy practices (Tanzi, 1995; Steinmo, 2003). Some have also tried to qualitatively assess the institutional influence of international organizations (Ring, 2010; Rixen, 2008; Lang, et al., 2012).

B. Tax Treaties and Model Treaties: Historical and Institutional Background

The historical background of the rise of tax treaties as an important instrument of international trade may help to explain the role and influence of institutional agents. Tax treaties were born out of the concern over double taxation as early as the late 19th century (Genschel & Rixen, 2015). At the time, growth in international trade introduced the potential of competing tax claims by multiple jurisdictions over the same streams of income. The institutional beginning of a coordinated international policy making on tax matters is commonly attributed to the International Chambers of Commerce (ICOC) (Graetz & O'Hear, 1997; Rixen, 2008). The ICOC banded together to represent international business interests following the First World War. During the early 1920s the ICOC urged the League of Nations to take measures to prevent double-taxation, which was seen as an impediment on reconstruction following the war. In response to ICOC's pressure, the League of Nations appointed a Committee of Experts to

develop principles for the prevention of double taxation. The most influential work of Committee of Experts was a 1923 report drafted by the four economists of the Committee: Edwin Seligman of the United States, Professor Luigi Einaudi of Italy, Professor G. W. J. Bruins of the Netherlands, and Sir Josiah Stamp of Great Britain (League of Nations, 1923).

Out of the 1923 report emerged "three great principles" of international taxation (Graetz & O'Hear, 1997): First, that "the classification and assignment of specific categories of income to source or residence should be determined by an objective test, 'economic allegiance'." Second, that "tax practices across the globe tended to underestimate the contribution of residence and to reflect a misguided belief in the naturalness and rightness of source-based taxation". Third, that progressivity in income taxes should be the prerogative of the residence jurisdiction. Some commentators suggested that the principles proposed by the 1923 report provide, to this day, the foundational features for the taxation of cross-border transactions (Ault, 1992; Avi-Yonah, 1996). Others dispute such a characterization (Graetz & O'Hear, 1997). The report clearly favored residence taxation (that is, the allocation of taxing rights to the country of residence of the investor), over source taxation (meaning, the allocation of taxing rights to the country in which the investment is located).

The 1923 report outlined economic principles, not legal language to implement them. The legal drafting task was assigned by the League of Nations to a committee of "Technical Experts" who produced a new report in 1925. "The 1925 Report was an effort to transform the pro-residence 1923 Report into a more balanced product" (Id.). The Technical Experts were then tasked with drafting model bilateral tax conventions. The purpose of drafting the model conventions was to achieve "a degree of uniformity between tax treaties by implementing bilateral tax treaties based on the Committee's draft convention." (Kobetsky, 2011). The

Technical Experts drafted several model treaties, each on a discrete topic in taxation. Several such models were drafted between 1925 and 1928. In 1928 the League of Nations held a general meeting to review, amend, and subsequently approve the conventions (League of Nations, 1928). The 1928 treaties eventually became the "definitive League model" (Graetz & O'Hear, 1997), and served as the basis of all subsequent models put forth by the League of Nations.

The 1928 meeting established a permanent Fiscal Committee in charge of the development of the League of Nations' models and to consider the allocation of international income between associate enterprises. Most importantly, the Fiscal Committee considered a 1933 report by Mitchell B. Carroll's (Carroll, 1933). Carroll's report surveyed the methods of apportionment used by countries in an attempt to distill general rules of profits allocation within a multinational enterprise. The most influential aspect of the Carroll report was the introduction of the arm's length principle in a newly published League of Nations model of 1935. Under the arm's length principle, affiliated companies must deal with each other at arm's length prices. The arm's length standard is considered the golden standard of intercompany pricing to this day (Picciotto, 2016). For example, in its most recent guidance on intercompany pricing, the OECD still describes the arm's length principle as "the standard that OECD member countries have agreed should be used for tax purposes by MNE groups and tax administrations." (OECD, 2017). The 1935 model also highlighted, for the first time, "U.S. leadership in international tax issues", as the 1935 model convention effectively adopted the U.S. transfer pricing rules then in place (Rixen, 2008).

Following the approval of the 1935 model, the League of Nations attempted to update the model on an annual basis until the beginning of World War II. With the war in Europe escalating, the League of Nations scaled down its affairs. However, before the League of

Nations' break in activity (and ultimate demise and replacement by the United Nations), the Fiscal Committee suggested the model treaty be redrafted, leading to a meeting at The Hague in 1940 to begin the process (League of Nations, 1946). Further meetings were held in Mexico City and in London, resulting in two new models: one in 1943 (The "Mexico Model") and one in 1946 (the "London Model"). Probably due to constraints on travel during World War II, the attendees in both drafting meeting varied significantly (Kobetsky, 2011). The 1943 Mexico meeting was attended mostly by Latin American participants, all capital-importing countries, who favored source-based taxation. The 1946 meeting had a robust attendance by capital export countries, who favored residence-based taxation. This resulted in significantly different models, with the Mexico Model leaning towards source taxation, while the London Model towards residence taxation. Many of the provisions and principles included in both the London and Mexico drafts can be found within the model treaties of the UN and the OECD today. The UN Model is seen as a successor to the Mexico Draft, while the OECD Model is a successor to the London Draft (Lennard, 2008). The two models represent clear historical institutional divergence in treaty development, marking a different tax policy path to capital exporting versus capital importing countries. Given the historical importance of the London and Mexico Models, we also explore their influence on tax treaty practices.

The United Nations did not seriously engage in model drafting again until the 1970s. The years following the World War II saw increased economic interdependence, particularly between countries of the Organisation for European Economic Co-operation (OEEC, the predecessor of the OECD). Responding again to pressure from the ICOC and from several member states, the OEEC established a Fiscal Committee that sought to draft a model convention in order to "effectively resolve the double taxation problems existing between OECD member countries"

(Vogel, Reimer, and Rust, 2014). The Fiscal Committee based its work on the London and Mexico models. Both models were understood to have significant influence on bilateral tax treaties adopted during the late 1940s and early 1950s, but none was viewed as a basis for a broad consensus. (Id.). The first OECD Model was published in 1963. The 1963 OECD Model seems to have aligned with the London Model, adopting a pro-residence stance favored by capital exporting countries. In the years following the adoption of the OECD Model, the OECD became "the main multilateral forum for international tax policy" (Rixen, 2008). In addition, new countries joining the OECD, particularly the United States, became influential. U.S. positions were fully incorporated into the drafting and revisions of the work of the OECD (Id.). The OECD since then updated its model regularly. Two new complete models were published in 1977 and in 1992. In 1991, however, the OECD recognized that the "revision of the Model Convention and the Commentaries had become an ongoing process" (OECD, 2014), and adopted a process under which the revisions to the model are published every two or three years.

While the OECD Model is not a binding document, it is generally accepted that the OECD Model "has had wide repercussions on the negotiation, application, and interpretation of tax conventions", even outside the OECD (Vogel, Reimer, and Rust, 2014; OECD, 2014). A summary report of a recent influential survey of the tax treaty practices in 37 countries, from both within and without the OECD (Lang, et al., 2012), concludes that "the influence the [OECD Model] on the general structure and clauses of bilateral tax treaties has gradually gained in importance so that it now affects those concluded with or even between non-OECD Member countries" (Pistone, 2012).

As the OECD was gaining prominence as the main institutional source for international tax policy, developing countries were dismayed by the emphasis on residence taxation. In

response, several attempts were made at drafting model conventions that shifted the focus to source taxation. The earliest attempt was probably the ANDEAN Model adopted by several Latin American countries in 1971 (Andean Community, 1971). This model is believed to have had relatively little influence (Rixen, 2008). The UN itself established a Committee of Experts on International Cooperation in Tax Matters in 1967. The work of this group of experts eventually led to the publication of the first UN Model in 1980. Unlike the OECD Model, the UN Model emphasized source taxation and the interests of developing countries. The UN Model is not regularly updated, though two major revisions have been published, one in 2001, and another in 2011 (United Nations, 2011). (An additional UN Model update was published in 2017, but as explained below, there is not enough recent data available to analyze it). The United Nations asserts in the preamble to the UN Model that both the OECD Model as well as the UN Model "have had a profound influence on international treaty practice" (Id.). Others dispute the characterization of the UN Model as influential, and argue that the UN Model's main influence is in negotiations between developed and developing countries, and that in any case such influence has gradually decreased over the course of the 20th century (Rixen, 2008; Pistone, 2012). In this paper we explore, among others, the institutional influence of both the OECD and UN models over time.

III. DATA AND METHODS

This section describes the methods for constructing statistical data from the text of international tax treaties.

A. Treaty Data

The corpus of treaties was downloaded from the online database maintained by the International Bureau of Fiscal Documentation (IBFD). The IBFD maintains the most expansive database of treaties on tax matters. We first download all tax treaties drafted in English as well as treaties for which an English-language translation is available. We then exclude non-income tax treaties, multilateral treaties, and treaties that have never entered into effect. Since only U.S., OECD and UN models are argued to have an influence on treaty drafting we exclude seven models published by institutions other than the OECD, UN or the United States. The resulting dataset is comprised of 4,502 bilateral treaties that at some point were in effect, and 23 model treaties. The earliest bilateral treaty in our dataset entered into force in 1942, and the latest is in 2015. There are 205 party-countries represented. We also separately add to our dataset the Mexico Model and the London Model, which we obtained from the UN website.

Each treaty has information on the parties, current status, conclusion date, effective date, and entry-into-force date. Figure 1 presents the number of bilateral tax treaties concluded by year. Figure 2 aggregates the total number of treaties in effect at any given year.

Figure 1

[Insert Figure 1 Here: Number of Treaties Concluded by Year]

Figure 2

[Insert Figure 2 Here: Active Treaties by Year]

Figure 1 and 2 demonstrate the increasing importance of bilateral tax treaties in international trade over time.

Finally, in order to assess whether legal language convergence is more apparent in certain areas than others, we take advantage of the fact that the IBFD database divides the text of the

treaties into 32 categories. The text of the treaties is thus split into sections, which allows for a section-by-section comparison. We manually divide the Mexico and London Models into categories that best match the categories in the IBFD database.

B. Constructing Text Features

We follow standard methods in the use of natural language processing in social science. A series of scripts reads through the text of the clauses and processes them as follows. First a clause is split into sentences. These sentences are then split into words. The vocabulary is then filtered such that any words not appearing in at least 10 clauses are excluded — these include some foreign-language terms, misspelled words, or place names, for example. Numbers are replaced with a special token, as are country names.

Next, the sentences are used to produce n-grams (phrases) up to a length of four words. These n-grams are filtered based on their parts of speech in order to obtain informative noun and verb phrases. The set of parts-of-speech sequences are based on Handler et al. (2016). The resulting text features include technical key phrases from international tax law such as "income from immovable property," "income from government securities," "preparatory or auxiliary character," "has an habitual abode," "through a permanent establishment," "fixed place of business," and "is the beneficial owner." These phrases provide much more legal information than single words, or n-grams regardless of the parts-of-speech. Moreover, this method captures the highly context-dependent meanings of individual words, such as "income" (e.g., "income from immovable property" versus "income from government securities"). Single words are included if they are nouns, adjectives, verbs, or adverbs, so uninformative stopwords like "not," by themselves, are excluded. But when they are part of an informative phrase, such as "are not included" and "are not residents," they are included.

15

In the final vocabulary of text features, words and phrases must occur in at least 10 clauses to be included. Single words are included if they are nouns, verbs, adjectives, or adverbs. In addition, party names and non-party countries are tagged with special tokens. The final vocabulary has 45,259 features. This includes 4,051 words, 13,469 bigrams, 18,447 trigrams, and 9,292 quadgrams. Using this vocabulary of features (words and phrases), we construct frequency distributions over features for each treaty in the corpus. The informative phrases are treated as single tokens and linked together, so if the word "course" appears as part of the phrase "in the normal course," it will not be included in the frequency distribution by itself. The outcome of this procedure is that each treaty is represented as a sparse vector of phrase frequencies. These vectors are used in the analysis. We explain the vector representation of words and documents at length in Online Appendix 2.

C. Computing Treaty Similarity

The treaty frequency vectors are then stacked into an *NxP* sparse matrix, where *N=4,052* treaties and *P=45,259* text features. We then compute the cosine similarity between each treaty vector. This vector of similarities, of length *N(N-1)=19,092,530* treaty pairs, gives the pair-wise similarity for each row in the matrix. The pair-wise similarities between treaties are used in the empirical analysis.

D. Metadata

For further checks and to flesh out our analysis, we collected a set of metadata to add to our treaties corpus.

First, we downloaded data from U.N. COMTRADE on the value of trade flows in each country for the year 2005. We matched the U.N. country identifiers to our data and computed the

share of world trade covered by our treaty corpus. We found that as of 2005, our treaties covered 19.6 percent of potential trading pairs (that is, potential links between all countries). However, when weighted by the value of trade, we see that our treaties cover 89.3 percent of world trade flows. This is because of the full set of potential trading partners, many countries do not trade with each other. Countries that don't trade very much do not need a tax treaty. Thus, any consensus identified in our research can be said to be relevant to the bulk of world trade.

Figure 3

[Insert Figure 3 Here: Active Treaties by Year, By Income Level of Parties]

Second, we categorized countries by current income status. We downloaded the World Bank classifications for high income, upper middle income, lower middle income, and low income. To simplify the analysis, we treated high and upper middle as high income, and we treated lower and lower middle as low income. We then categorized each trading pair has high-high, high-low, or low-low. Figure 3 shows the distribution of treaty parties by current income status over the time period of our data set.

IV. RESULTS

A. Convergence in General

This section provides evidence on whether tax treaties are converging in the similarity of their language. To assess overall similarity in tax treaty language over time we measure the pair-wise similarity of each possible pair of treaties in force at any given year, and calculate the mean similarity in each year. Figure 4 presents these statistics. The dashed error spikes provide the 25th and 75th quantiles of these measures by year.

Figure 4

[Insert Figure 4 Here: Pair-wise Text Similarity of Treaties by Year]

Figure 4 shows a clear increase in similarity from around 1970 to around 2010. In the top panel, we look at the full set of active treaties. In the bottom panel, we only include the similarity between the new treaties entered into force in a year. We can see that there is some decrease in the similarity of new treaties, since 2010.

The trend holds when we exclude treaties to which at least one country member is an OECD member.

The trend towards convergence becomes clear in the 1970s, in the time when the OECD was the only international institution actively engaged in tax treaty policy. The UN did not publish a model treaty until 1980, and the United States published its first model in 1976. The fact that the OECD dominance on international tax matters was not challenged may have been a contribution factor to prolonged OECD influence (this is further discussed below).

We next turn to measure the pair-wise similarity among each category, applying the same method we use to assess overall convergence. Table 1 shows absolute level of legal language similarity for each of the 32 categories. Absolute similarity is shown for both 1965 and 2015. We also present the change in similarity over such period to show convergence (divergence) within each category, both in absolute and proportional terms. The table is ordered from to most to least similar categories per our 2015 measurements.

Table 1

[Insert Table 1 here: Pairwise Similarity of Specific Categories in Active Treaties]

Note that there is convergence across a range of categories. However, some categories are not becoming more similar, or even diverging. This suggests that there are particular legal requirements becoming more similar, and that the effects are not driven just by increasing standardization of legal language generally. If our results were due to all language becoming more similar, we would see equal increases in similarity across categories. We mark in the table

areas that have experienced significant convergence over the tested period (props sim above 0.5, marked with "+"; or above 1.0 marked with "++"), or divergence (prop sim below zero, marked with "-").

Next, we look at convergence by country income classification. In Figures 5, 6, and 7, we show these convergences. We see that treaty language is converging for all of these classes.

Figure 5

[Figure 5 Pair-wise Text Similarity of Active Treaties By Year, with Two High-Income Parties]

Figure 6

[Figure 6 Pair-wise Text Similarity of Active Treaties By Year, One High and One Low-Income

Party]

Figure 7

[Figure 7 Pair-wise Text Similarity of Active Treaties by Year, Two Low-Income Parties]

V. WHICH MODEL TREATIES ARE THE MOST INFLUENTIAL?

This section provides evidence on which model treaties have the largest impact (in terms of text content) on bilateral treaties.

Figure 8

[Figure 8: Average Similarity of Treaties in Force to Model Treaties]

Figure 8 shows the trends in similarity of active treaties to the three models, OECD, United Nations, and United States. It is possible to identify discrete jumps when the various actors introduce new models. A discrete jump upwards means that the new model is more similar to the existing stock of treaties. A discrete jump downwards means that the new model is less similar to the existing stock of treaties. On average, recent active treaties are most similar to OECD and United Nations Treaties. Interestingly, in recent years active treaties seem to be

19

slightly more similar to the latest UN Model (though the difference of similarity to the OECD Model is negligible). As we discuss further below, the introduction of new UN Models seems to have little short-term effect on treaties adopted following the introduction of the model. On the other hand, it seems that existing treaty practices, as well as the OECD Model, very slowly converge towards UN legal language. This possibly suggests a very slow process of abandoning favoritism for residence-based taxation, in favor of source-based taxation, even among OECD countries.

Figure 9

[Figure 9: Similarity of New Treaties to OECD Models]

Figure 10

[Figure 10: Similarity of New Treaties to UN Models]

Figure 11

[Figure 11: Similarity of New Treaties to US Models]

In Figures 9, 10, and 11, we look at new treaty similarity to the collection of models in our comparison corpus. The 1963 OECD Model was most influential initially, but then the 1977-1998 models have been most influential. Treaties are quite consistent in their similarity to the UN Models. The U.S. Models are all quite different from each other, especially the 2016 U.S. Model.

We next turn to question whether the introduction of a new model treaty has an observable effect on actual treaty drafting in the short-to-medium term.

Figure 12

[Figure 12 - Relative Similarity to New Models Relative to Old Models, By Conclusion Year]

[12a. New OECD Models]

[12b. New UN Models]

[12c. New U.S. Models]

Figure 12 plots the relative similarity of newly concluded treaties to the newest model, relative to the previous model. Formally, this is the average cosine similarity of treaties concluded in a year to the new model, divided by the average cosine similarity of those treaties to the old model. The figure plots the average of this measure in ten months before and after a new model is introduced, separately for OECD, United Nations, and U.S. models. An increase in the measure after the treaty means that new treaties are following the new model, in the sense that they are more similar to the new model. An upward-sloping pre-trend would mean that the new model is responding to pre-existing trends in tax treaty language.

In the context of new OECD Models, the response is a sloped upwards, post-model introduction. This suggests a convergence over time of newly concluded treaties towards new OECD Models. It is therefore reasonable to conclude that the introduction of OECD Models has an effect on new treaty drafting.

In contrast, we see no significant trend before or after new models for the U.N. Models or the U.S. Models. In the context of new U.S. Models, one notices an upward-sloping pre-trend. This suggests that the new model is responding to pre-existing trends in treaty changes. Indeed, some commentators have suggested that the U.S. Model is sometimes revised to conform with existing OECD practices. (Brauner & Marian, 2014).

Table 2

[Table 2: Change in Similarity]

To round out this analysis, we looked at the change in similarity computed for each new model. Table 2 gives, for each new model, the similarity of its new model, minus the similarity of the previous model, for the tax treaties enacted in the interim. A positive number means that

the model is "following" new innovations in the treaties. But we see they are almost all negative, meaning that models try to break new ground — they are leaders rather than followers.

Next, we look at the relative similarity of treaties to each model, relative to the previous model, separately for each model. This is a monthly, rather than annual, analysis, so we can look at models that were released somewhat close to each other. We can use the metric to identify the influential models, as well as unpopular models. Influential models include the OECD's 1977, 1992, 2005, 2008, 2010, and 2014 models, and the U.S. 1996 Model. Models that caused a backlash include the OECD 2003 model and the U.S. 2006 Model.

Figure 13

[Figure 13: Influential Models]

Figure 14

[Figure 14: Unpopular Models]

Next, we looked at the simple question of whether joining the OECD has an impact on similarity to the OECD Model. Since each treaty has two parties, there are actually two potential experiments here: one party joining, versus a second party joining. In Figure 15 we show these effects in an event study framework. We can see that one party joining the OECD doesn't have much of an effect on similarity to the OECD Model. The second party joining actually seems to have a negative effect. This suggests that joining the OECD, by itself, is not a major factor driving the influence of that model.

Figure 15

[Figure 15 - Effect of Joining OECD on Treaty Similarity to OECD Model]

Finally, we are interested in similarity to the London and Mexico models. This was done using the same method as for the other models. The similarity over time is reported in Figure 16. We see that early on, both models were equally similar to treaty text. But since the 1970s, the

Mexico Model is more similar. We found a similar trend when breaking up by the income classes of the treaty parties.

Figure 16

[Figure 16 - Text Similarity to London Model and Mexico Model]

VI. CONCLUSION AND DISCUSSION

Our empirical investigation has important implications to the international tax regime debate. Specifically, if countries are free to adopt whatever tax rules they wish, a high level of variance in tax treaty language is expected. The reason is that in tax treaty negotiations, countries will try to adopt the position that best serves their national interest (Dagan, 2000). Each pair of countries presents a different set of negotiating circumstances. For example, one country may be a net capital exporter in relation to one treaty partner, but a capital importer in relation to another. A country may hold a strong negotiating position vis-a-vis one treaty partner (for example, due to economic size), but a weak stance against another. Different pairs of countries may present varying levels of kinship or animosity, whether diplomatic or cultural. Given the varied sets of circumstances applicable to each particular treaty, it is reasonable to expect a high level a variance among treaties. Our findings, however, point to convergence in legal language, which may suggest that countries are guided by transnational legal considerations.

We find that convergence in legal language is most clearly observed in the context of intercompany pricing, taxation of cross-border business income, and in the context of mutual agreement procedures. The lowest levels of convergence are observed in connection with certain definitional issues (such as the taxes and the geographical extent to which treaties apply), on the question on how to relieve double taxation, as well as in the context of assistance in collection of taxes.

23

We also explore the institutional aspect of consensus building in tax treaties. We find the OECD Model to be the most influential model. In the years following the adoption of a new OECD Model there is a clear trend of convergence in newly adopted bilateral tax treaties towards the language of the new OECD Model. This suggests that the OECD plays an important role in facilitating international legal consensus on tax matters through the publication of its model treaty.

We also find that model treaties published by the UN have little observable effect in the shot-to-medium term. However, current treaty practices seem to align themselves with the UN Model of 2011 more than with the OECD Model. It is therefore reasonable to accept an argument according to which the UN tax policies may have a long-term effect, representing a slow shift from residence to source-based taxation, even among developed countries.

Overall, we believe our findings support the argument that a trend towards international legal consensus exists, at least on certain matters, and that the OECD is the institutional source of the consensus building process. The OECD seems to play an effective role as a quasi-formal international tax organization on tax treaty matters. We stop short of concluding that a customary international law of taxation exists, though we believe our findings lend some support to such arguments.

REFERENCES

Andean Community, 1971. *Income and Capital Model Tax Treaty (Unofficial Translation),* s.l.: s.n.

Ault, Hugh J., 1992. "Corporate Integration, Tax Treaties, and the Division of International Tax Base: Principles and Practices." *Tax Law Review* 47, 565–608.

Ault, Hugh J., 2009. "Reflections on the Role of the OECD in Developing International Tax Norms." *Brooklyn Journal of International Law* 34 (3), 757–781.

Avi-Yonah, Reuven S., 1996. "The Structure of International Taxation: A Proposal for Simplification." *Texas Law Review* 74, 1301–1359.

Avi-Yonah, Reuven S., 2000. "Commentary." *Tax Law Review* 53 (2), 167–175.

Avi-Yonah, Reuven S., 2004. "International Tax as International Law." *Tax Law Review* 57 (4), 483–502.

Avi-Yonah, Reuven S., 2007. *International Tax as International Law: An Analysis of the International Tax Regime.* Cambridge University Press, New York, NY.

Avi-Yonah, Reuven S., 2009. "Double Tax Treaties: An Introduction." In Sauvant, Karl P., and Lisa E. Sachs (eds.), *The Effect of Treaties on Foreign Direct Investment: Bilateral Investment Treaties, Double Taxation Treaties, and Investment Flows,* 99–106. Oxford University Press, New York, NY.

Brauner, Yariv, 2003. "An International Tax Regime in Crystallization." *Tax Law Review* 56, 259–328.

Brauner, Yariv, 2005. "International Trade and Tax Agreements may be Coordinated, but not Reconciled." *Virginia Tax Review* 25, 251–311.

Brauner, Yariv, 2016. "Treaties in the Aftermath of BEPS." *Brooklyn Journal of International Law* 41 (3).

Carroll, Mitchell B., 1933. *Report on Methods of Allocating Taxable Income,* s.l.: s.n.

Christians, Allison, 2007. "Hard Law, Soft Law, and International Taxation." *Wisconsin International Law Journal* 25, 325–333.

Cockfield, Arthur J., 2006. "The Rise of the OECD as Informal "World Tax Organisation" Through National Response to E-Commerce Tax Challanges." *Yale Journal of Law and Technology* 8, 136–187.

Dagan, Tsilly, 2000. "The Tax Treaties Myth." *New York University Journal of International Law & Politics* 32, 939–996.

Genschel, Phillip and Thomas Rixen, 2015. "Settling and Unsettling the Transnational Legal Order of International Taxation." In Halliday, Terence C. and Shaffer, Gregory (eds.), *Transnational Legal Orders,* 154–183. Cambridge University Press, New York, NY.

Graetz, Michael J., and Michael M. O'Hear, 1997. "The "Original Intent" of U.S. International Taxation." *Duke Law Journal* 46, 1021–1109.

Handler, Abram, et al, 2016. "Bag of What? Simple Noun Phrase Extraction for Text Analysis." *Proceedings of the First Workshop on NLP and Computational Social Science 2016*

Kane, Mitchell, 2004. "Strategy and Cooperation in National Responses to International Tax Arbitrage." *Emory Law Journal* 53 89–169.

Kaufman, Nancy H., 1998. "Fairness and the Taxation of International Income." *Law and Policy in International Business* 29, 145–203.

Kobetsky, Michael, 2011. *International Taxation of Permanent Establishments: Principles and Policy.* s.l.:s.n.

Lang, Michael, Pasquale Pistone, Josef Schuch and Claus Staringer, 2012. *The Impact of the OECD and UN Model Conventions on Bilateral Tax Treaties.* Cambridge University Press, Cambridge, UK.

League of Nations, 1923. *League of Nations, Economic and Financial Commission, Report on Double Taxation Submitted to the Financial Committee.* League of Nations, Geneva.

League of Nations, 1927. *Report Presented by the Committee of Technical Experts on Double Taxation and Tax Evasion, Double Taxation and Tax Evasion,* s.l.: s.n.

League of Nations, 1928. *Report Presented by the General Meeting of Governmental Experts on Double Taxation and Tax Evasion, Double Taxation and Tax Evasion,* s.l.: s.n.

League of Nations, 1946. *Fiscal Committee, Report on the Work of the Tenth Session of the Committee,* s.l.: s.n.

Lennard, Michael, 2008. "The Purpose and Current Status of the United Nations Tax Work." *Asia-Pacific Tax Bulletin,* 23–20.

OECD, 2014. *Model Tax Convention on Income and on Capital 2014 (Full Version).* s.l.:s.n.

OECD, 2016. *Multilateral Convention to Implement Tax Treaty Related Measures to Prevent Base Erosion and Profit Shifting.* OECD, Paris.

OECD, 2017. *OECD Transfer Pricing Guidelines for Multinational Corporations and Tax Administrators.* OECD, Paris.

OECD, 2017. *Revenue Statistics.* Paris: OECD.

Organisation for Economic Cooperation and Development, 2015. *OECD/G20 Base Erosion and Profit Shifting Project: Explanatory Statement.* s.l.:s.n.

Organisation for Economic Cooperation and Development, 2010. *Transfer Pricing Guidelines for Multinational Enterprises and Tax Administrations.* s.l.: Organisation for Economic Cooperation and Development.

Picciotto, Sol, 2016. "Taxing Multinational Enterprises as Unitary Entities." *Tax Notes International* 82, 895–916.

Pistone, Pasquale, 2012. "General Report." In Lang, Michael and Pasquale Pistone (eds.) *The Impact of the OECD and UN Model Conventions on Bilateral Tax Treaties.* s.l.:Cambridge University Press, 1–37.

Ring, Diane, 2010. "Who is Making International Tax Policy? International Organizations as Power Players in a High Stakes World." *Fordham International Law Journal* 33, 649–722.

Rixen, Thomas, 2008. *The Political Economy of International Tax Governance.* Palgrave MacMillan, New York, NY.

Roin, Julie, 2002. "Taxation without Coordination." *Journal of Legal Studies* 31, 61–94.

Rosenbloom, H. David, 2000. "International Tax Arbitrage and the "International Tax System"." *Tax Law Review* 53, 137–166.

Seligman, Edwin R. A., 1928. *Double Taxation and International Fiscal Cooperation.* s.l.:s.n.

Slemrod, Joel, 2004. "Are Corporate Tax Rates, or Countries, Converging?" *Journal of Public Economics* 88(6), 1169–1186.

Steinmo, Sven, 2003. "The Evolution of Policy Ideas: Tax Policy in the 20th Century." *British Journal of Politics and International Relations* 5(2), 206–236.

Swank, Duane and Sven Steinmo, 2002. "The New Political Economy of Taxation in Advanced Capitalist Democracies." *American Journal of Political Science* 46(3), 642–655.

Tanzi, Vito, 1995. *Taxation in an Integrating World.* Brookings Institution, Washington DC.

Tanzi, Vito, 1999. "Is there a Need for a World Tax Organization?" In Raziv, Assaf and Efraim Sadka (eds.) *The Economics of Globalization,* 173–186. Cambridge University Press, New York, NY.

United Nations, 2011. *Model Double Taxation Convention Between Developed and Developing Countries.* United Nations, New York, New York.

U.S. Department of Treasury, 2016. *White Paper, The European Commission's Recent State Aid Investigations on Transfer Pricing Rulings.* s.l.:s.n.

Vega, Alberto, 2012. "International Governance through Soft Law: The Case of the OECD Transfer Pricing Guidelines." *Max Planck Institute for Tax Law and Public Finance Working Paper* 2012–05.

Vogel, Klaus, Ekkehart Reimer and Alexander Rust, 2014. *Klaus Vogel on Double Taxation Conventions.* s.l.:s.n

Figure 1

Number of Treaties Concluded by Year
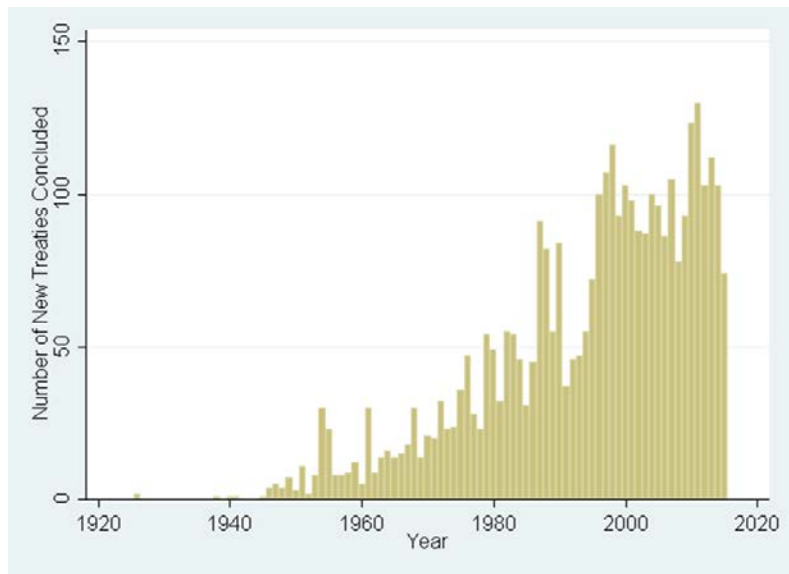
Figure 2

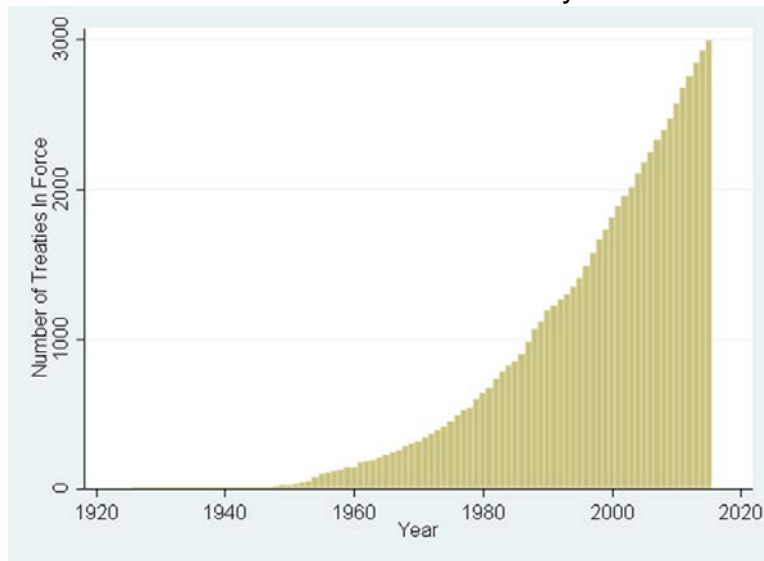Number of Treaties in Force by Year

Figure 3

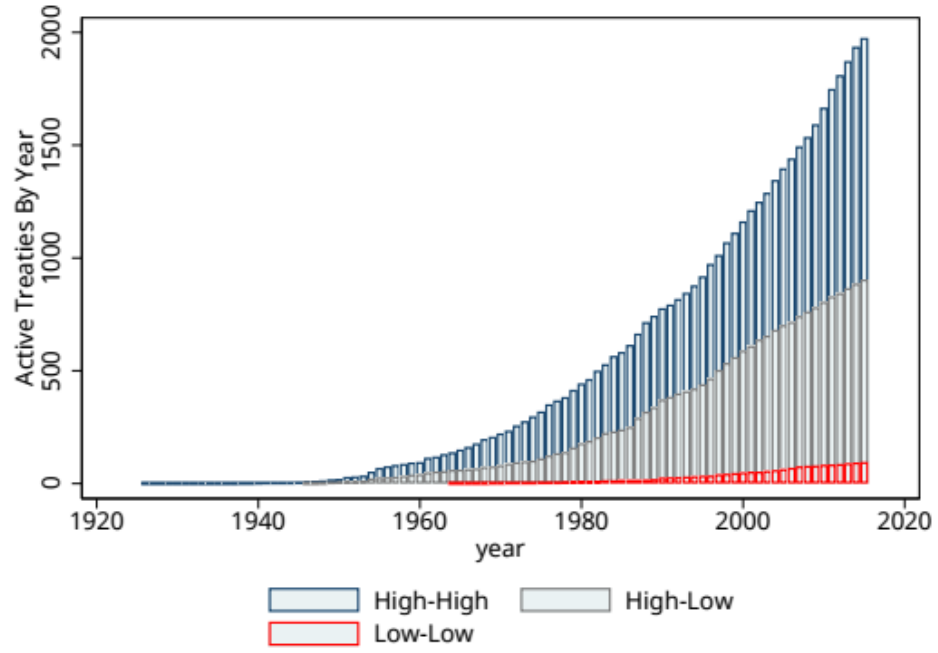Active Treaties by Year, By Income Classification of Parties

Figure 4

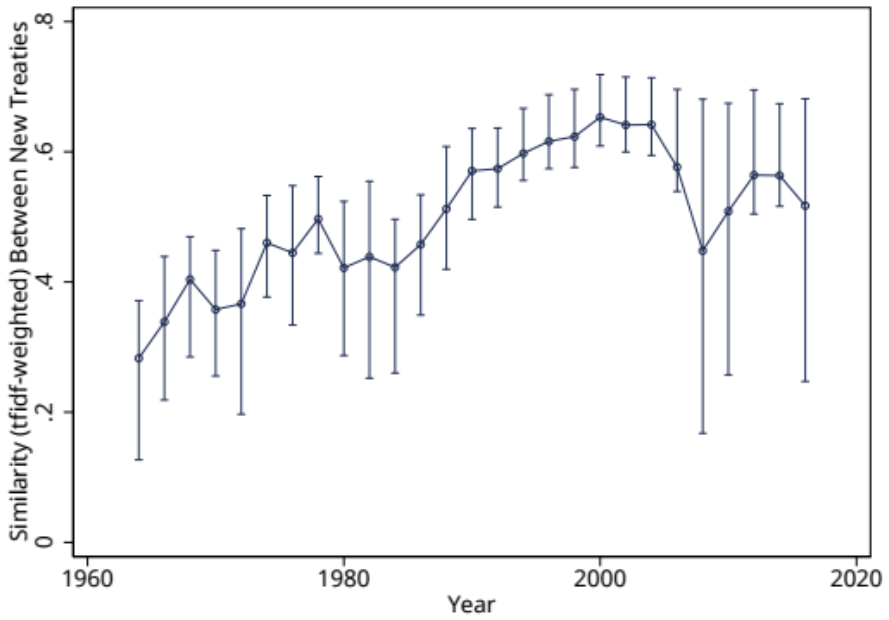Pair-wise Text Similarity of Treaties by Year

Table 1

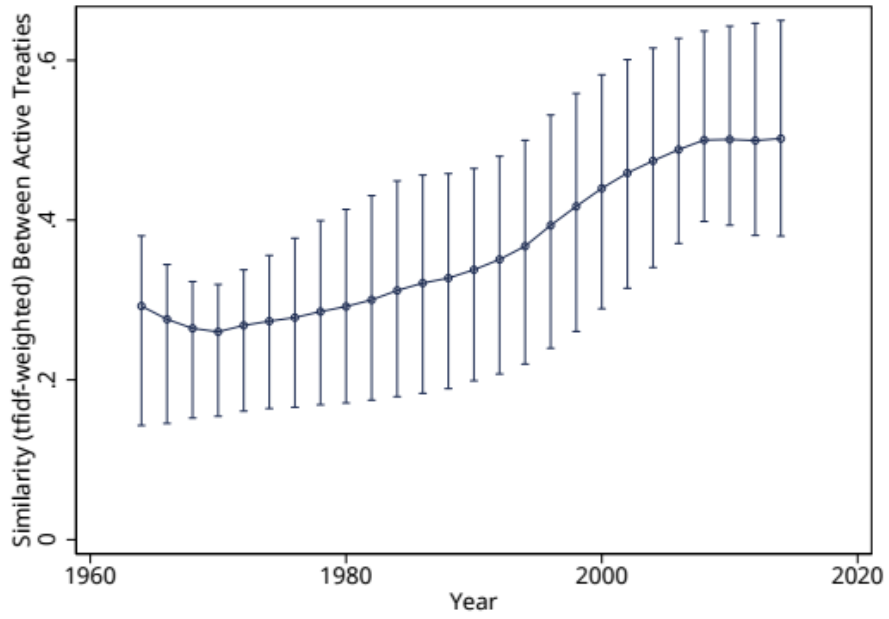Pairwise Similarity of Specific Categories in Active Treaties

Category	1965 sim	2015 sim	Change	Prop change	Convergence
Associated Enterprises	0.794	0.859	0.065	0.082	
Permanent Establishment	0.552	0.829	0.277	0.501	
Director Fees and Remuneration of Top Officials	0.330	0.776	0.446	1.353	++
Business Profits	0.562	0.765	0.203	0.361	
Income from Employment	0.433	0.750	0.317	0.732	+
Mutual Agreement Procedure	0.261	0.744	0.482	1.848	++
Capital	0.250	0.733	0.483	1.934	++
Dividends	0.350	0.720	0.370	1.055	++
Non-Discrimination	0.551	0.719	0.168	0.306	
Residence	0.364	0.719	0.355	0.973	+
Income from Royalties	0.242	0.714	0.472	1.951	++
Income from Immovable Property	0.189	0.695	0.507	2.682	++
Persons Covered	0.322	0.677	0.355	1.103	++
Exchange of Information	0.496	0.666	0.170	0.342	
Interest	0.251	0.665	0.413	1.643	++
Capital Gains	0.266	0.645	0.379	1.423	++
Independent Personal Services	0.330	0.642	0.313	0.948	+
Government Service	0.302	0.625	0.324	1.071	++
Entertainers and Sportspersons	0.464	0.612	0.148	0.320	
Other Income	0.301	0.608	0.307	1.019	++
Members of Diplomatic Missions	0.222	0.554	0.333	1.503	++
General Definitions	0.518	0.515	-0.003	-0.005	-
Territorial Extension	0.585	0.459	-0.127	-0.217	-
Pensions	0.386	0.446	0.060	0.154	
Taxes Covered	0.309	0.435	0.127	0.411	
Double Taxation (Exemption Method)	0.230	0.401	0.171	0.744	+
Shipping, Waterways Transport, and Air Transport	0.280	0.387	0.107	0.380	
Students	0.450	0.381	-0.069	-0.154	-
Termination	0.383	0.339	-0.043	-0.113	-
Assistance in the Collection of Taxes	0.177	0.288	0.111	0.627	
Double Taxation (Credit Method)	0.152	0.275	0.123	0.812	
Entry Into Force and Implementation	0.345	0.260	-0.085	-0.248	-

Figure 5

Pair-wise Text Similarity of Active Treaties by Year, with Two High-Income Parties
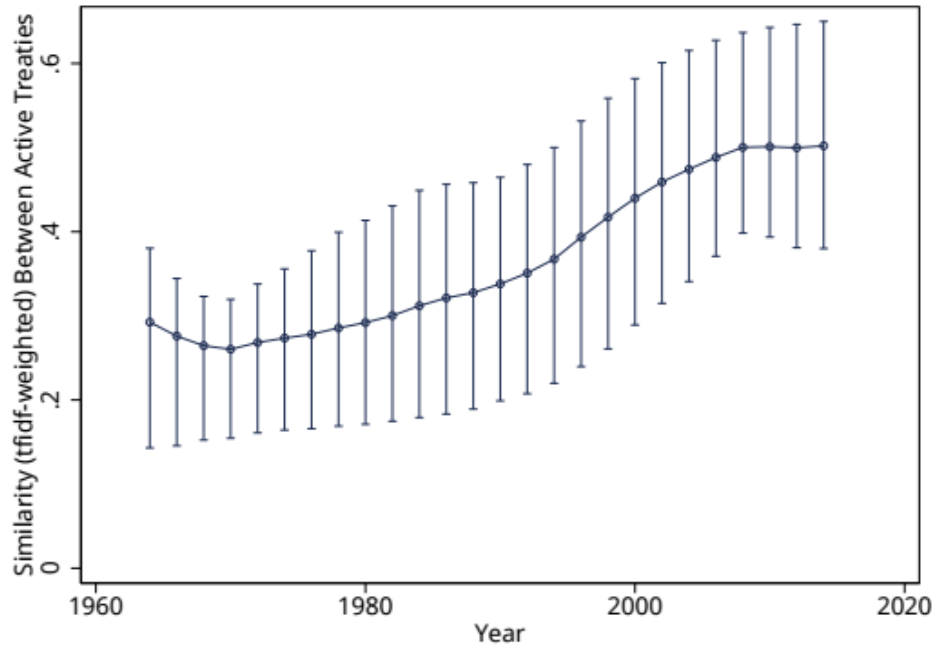
Figure 6

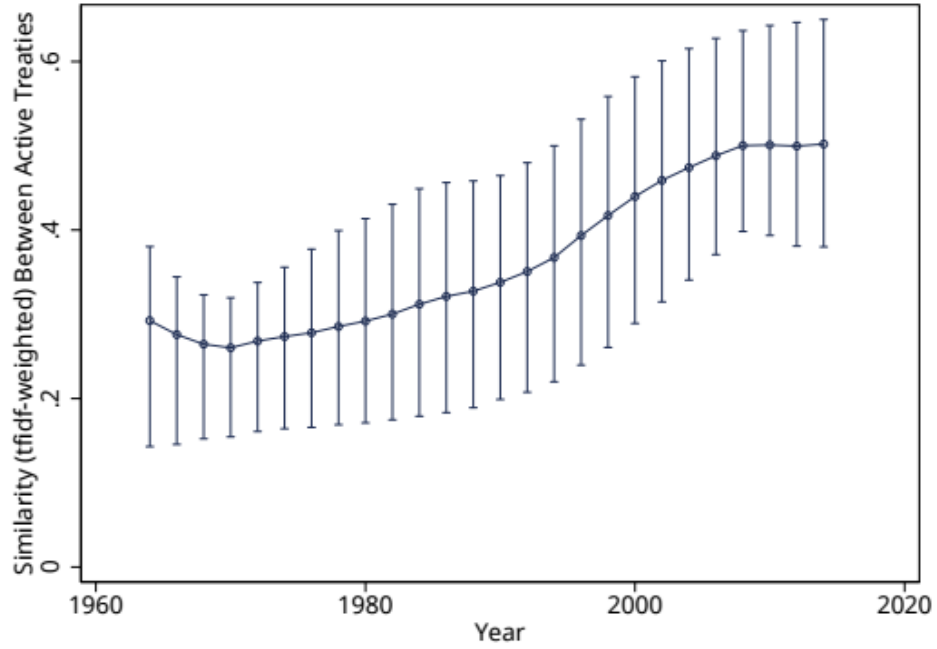Pair-wise Text Similarity of Active Treaties by Year, One High and One Low-Income Party

Figure 7

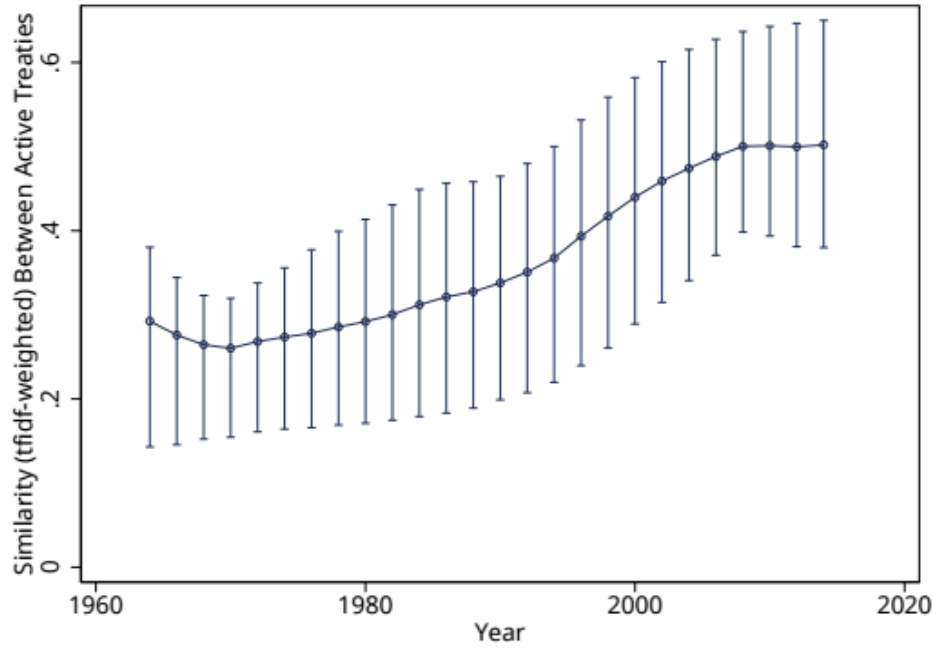Pair-wise Text Similarity of Active Treaties by Year, Two Low-Income Parties

Figure 8
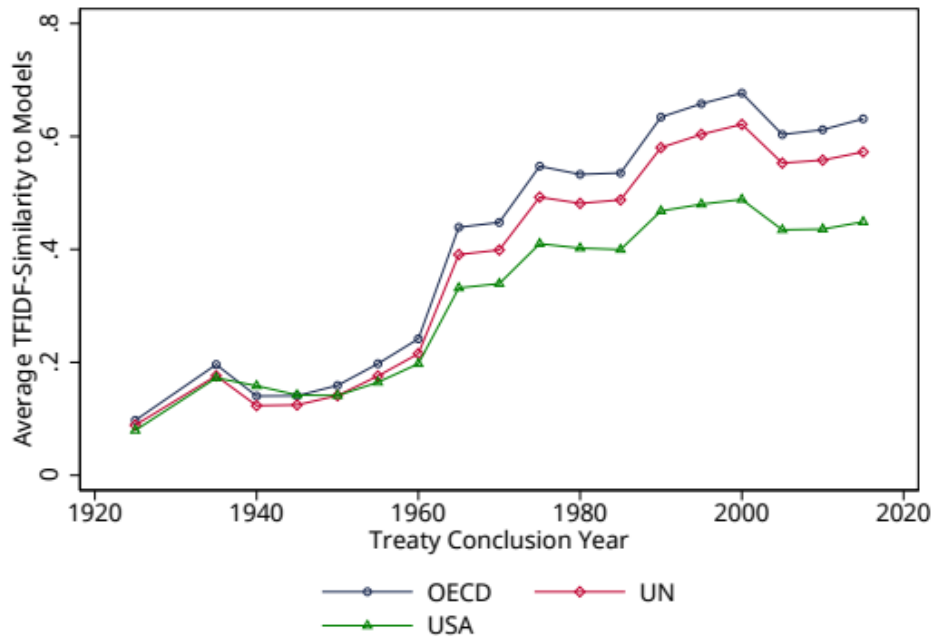
Average Similarity of Treaties in Force to Model Treaties
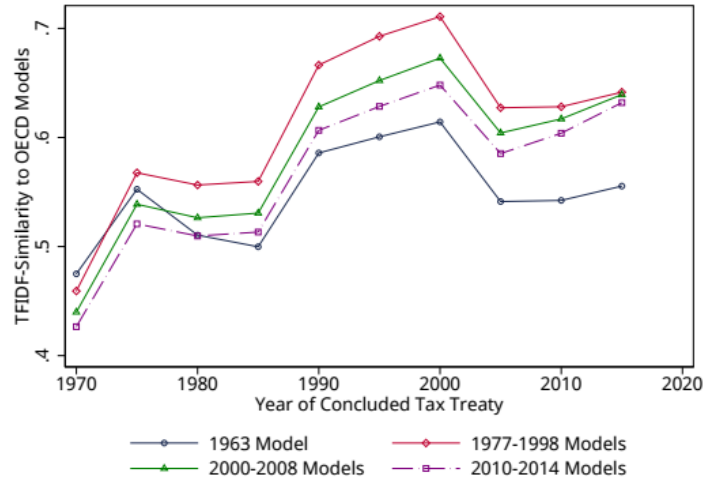
Figure 9

Similarity of New Treaties to OECD Models
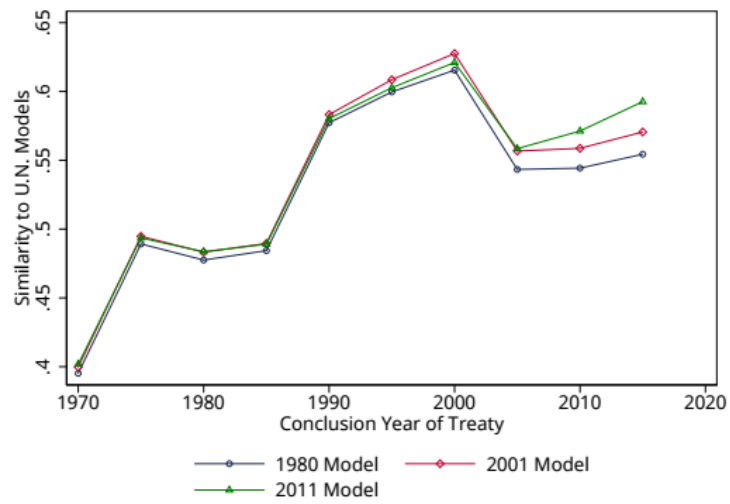
Figure 10

Similarity of New Treaties to UN Models

Figure 11

Similarity of New Treaties to U.S. Models
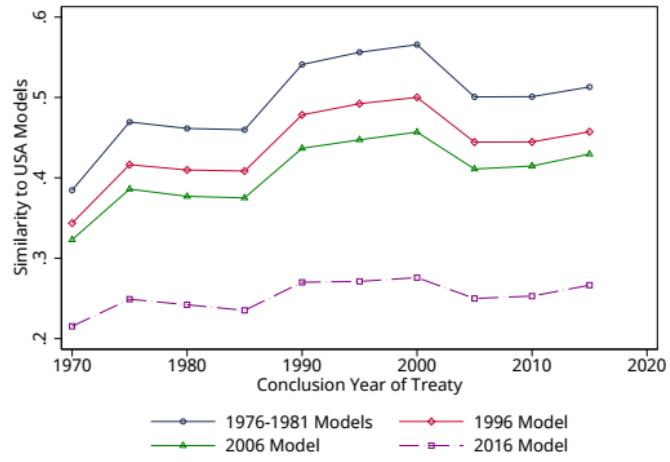
Figure 12

Relative Similarity to New Models Relative to Old Models, By Conclusion Year

12a. New OECD Models
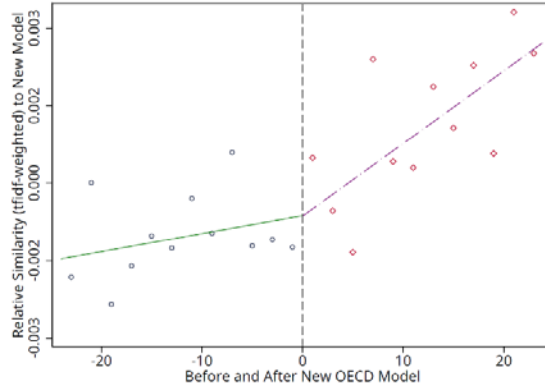
12b. New UN Models
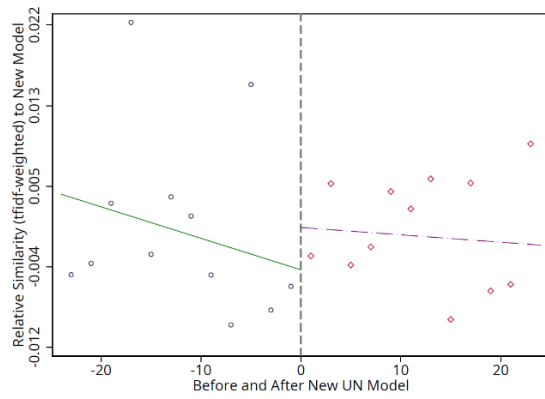
12c. New U.S. Models
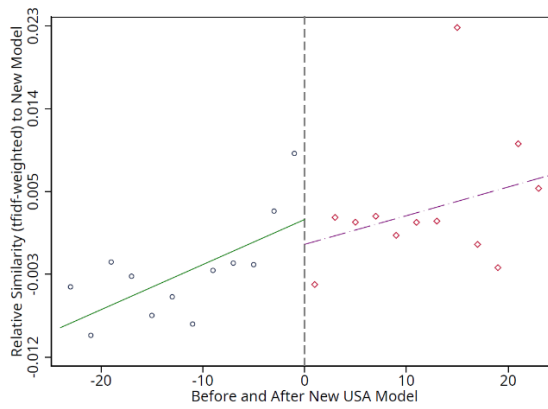
Table 2

Change in Similarity After Each Model

Model & Year	Change in Similarity
OECD 1977	-0.0263
OECD 1992	-0.0007
OECD 1996	-0.0014
OECD 1998	-0.0001
OECD 2000	-0.0076
OECD 2003	-0.0055
OECD 2005	-0.0016
OECD 2008	-0.0014
OECD 2010	-0.0043
OECD 2014	-0.0018
UN 2001	-0.0014
UN 2011	-0.0004
USA 1977	-0.0008
USA 1981	-0.0034
USA 1996	-0.0211
USA 2006	-0.0137
USA 2016	-0.1491
Average	-0.0198

Figure 13

Influential Models
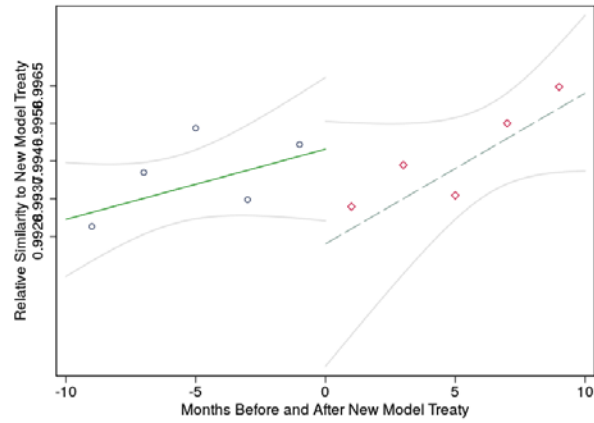
OECD, 1977
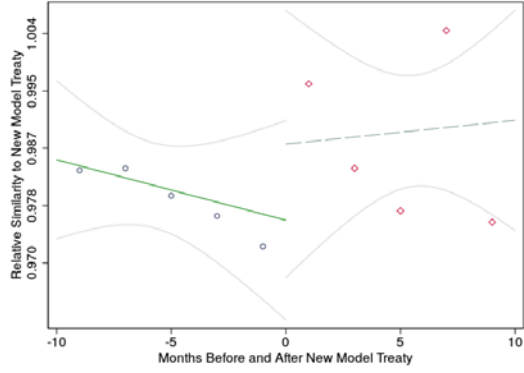
OECD, 1992
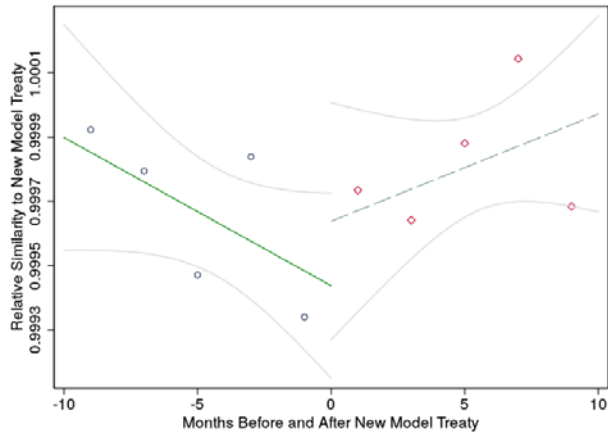
Figure 14

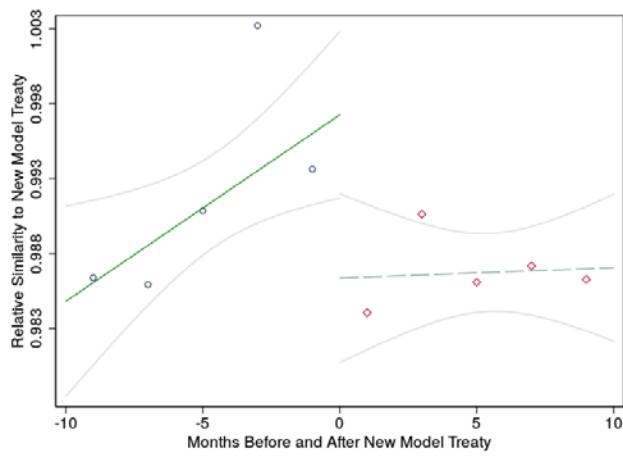Unpopular Models

OECD, 2002

United States, 2006

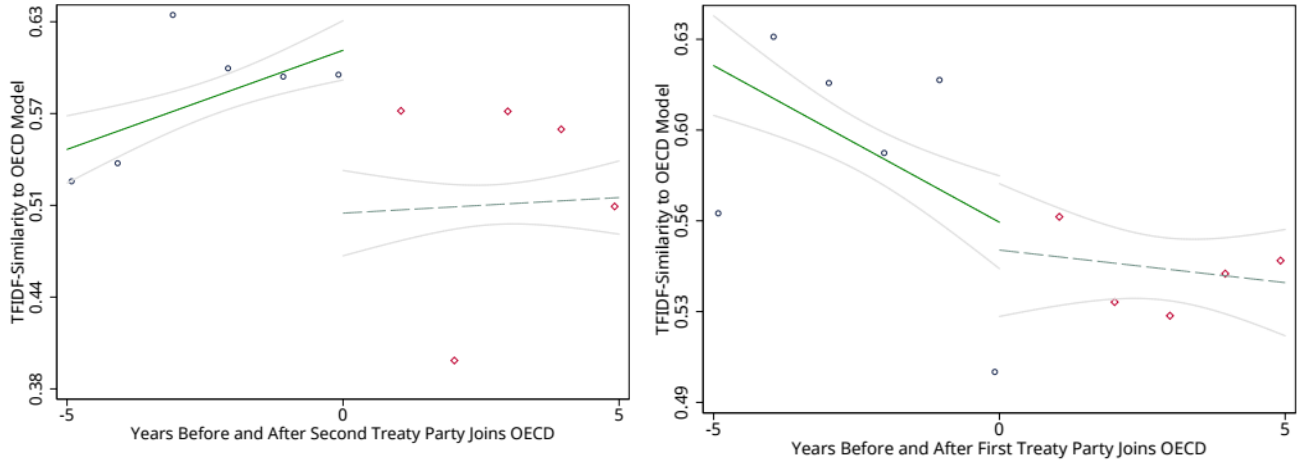**Figure 15**

Effect of Joining OECD on Treaty Similarity to OECD Model

Figure 16

Text Similarity to London Model and Mexico Model