doi:10.1017/ajil.2018.103

Giving the Treaty a Purpose: Comparing the Durability of Treaties and Executive Agreements

*By Julian Nyarko**

Abstract

Scholars have argued that Senate-approved treaties are becoming increasingly irrelevant in the United States, because their role can be fulfilled by their close but less politically costly cousin, the congressional-executive agreement. This study demonstrates that treaties are more durable than congressional-executive agreements, supporting the view that there are qualitative differences between the two instruments. Abandoning the treaty may therefore lead to unintended consequences by decreasing the tools that the executive has available to design optimal agreements.

I. Introduction

The United States is an international anomaly in that it has two largely interchangeable commitment devices to conclude the vast majority of its agreements with other states.[1] One instrument is the treaty. Treaties follow the advice and consent procedure set forth in Article II of the Constitution, which requires a two-thirds majority in the Senate in order for a treaty to be ratified and to become binding international law.[2] In place of the treaty, commitments can also be made in the form of a congressional-executive agreement, which requires only a simple majority in both the House of Representatives and the Senate.

Many scholars are skeptical of the utility of having two seemingly similar policy instruments to conclude international agreements, with most of the critique directed at the treaty for its supposed inflexibility and irrelevance. To be sure, criticism of the treaty is nothing new and dates back to at least the 1940s.[3] During that time, the discussion revolved around the

* Postdoctoral Fellow in Empirical Law and Economics, Ira M. Millstein Center for Global Markets and Corporate Ownership, Columbia Law School. For very helpful comments and suggestions, I am grateful to Andrew Guzman, Katerina Linos, Bertrall Ross, Jean Galbraith, John Yoo, Kevin Quinn, Robert Powell, Robert Cooter, Beth Simmons, William-Burke White, as well as the participants of the 2018 Perry World House at Penn Workshop on International Law, Organization, and Politics for Junior Scholars and Advanced Graduate Students.

[1] Oona A. Hathaway, *Treaties' End: The Past, Present, and Future of International Lawmaking in the United States*, 117 Yale L.J. 1236, 1239 (2008) (pointing out that "virtually no other country" has a two-track procedure of making international law like the United States does).

[2] U.S. Const. Art. II, § 2, cl. 2.

[3] Wallace McClure, International Executive Agreements: Democratic Procedure Under the Constitution of the United States 378 (1941) (arguing that the treaty should be replaced by the executive agreement, safe for the exception where "no public opinion exists and no question as to [the treaties'] acceptability

question of whether it is permissible to use both instruments interchangeably, with a particular focus on the constitutional limits on substituting congressional-executive agreements for treaties. However, over time it became clear that neither courts nor the State Department showed much concern with delineating constitutional limits on the interchangeability of the two commitment devices.[4] Take, for instance, former U.S. State Department Legal Adviser Harold Koh, who suggests that there are only two reasons why the State Department uses treaties, namely comity toward Congress and the "powerful political message" that is sent to the world through the treaty ratification process. With respect to the interchangeability of the instruments, Koh considers it the "long-dominant" view that it is constitutionally permissible to use congressional-executive agreements in place of the treaty.[5]

Today, the debate around interchangeability has resurfaced, albeit in a different form. The contemporary focus lies not on the question of whether it is permissible to use both instruments interchangeably, but on whether treaties and congressional-executive agreements can be used interchangeably as a matter of policy. During the Obama administration, only twenty treaties were approved by the Senate, the lowest number of approvals during a presidential term since President Ford.[6] At the same time, the popularity of the congressional-executive agreement seems unwavering, with several hundred international agreements having been concluded as congressional-executive agreements in the same time span.[7] Faced with this empirical reality, it has been contended that the treaty does not serve much practical purpose today, with some commentators even going as far as suggesting that the treaty should be abandoned altogether.[8]

Why, the argument goes, should presidents go through the slow and cumbersome advice and consent procedure of the treaty if their policy objectives can be fulfilled more easily by use of congressional-executive agreements, which are not similarly constrained?[9] After all, the latter's authorization can be granted broadly and *ex ante* through simple majoritarian approval, thus allowing the president to conclude a myriad of agreements authorized under a single congressional act.[10] If we see treaties used today, this account suggests, it would be for reasons that are orthogonal to the quality of the instrument itself, such as historical convention or selective senatorial preferences.[11]

arises"); Edwin Borchard, *Book Review: International Executive Agreements: Democratic Procedure Under the Constitution of the United States*, 42 COLUM. L. REV. 887 (1942) (rebutting McClure's argument, characterizing it as unconstitutional); *see also* Edwin Borchard, *Shall the Executive Agreement Replace the Treaty?*, 53 YALE L.J. 664 (1944) (characterizing executive agreements as the weaker commitment device).

[4] *See, e.g.*, United States v. Belmont, 301 U.S. 324 (1937); United States v. Pink, 315 U.S. 203 (1942); Dames & Moore v. Regan, 453 U.S. 654 (1981).

[5] Harold H. Koh, *Treaties and Agreements as Part of Twenty-First Century International Lawmaking*, *in* DIGEST OF UNITED STATES PRACTICE IN INTERNATIONAL LAW 91, 91–92 (CarrieLyn D. Guymon ed., 2012).

[6] This number is based on a count of treaty documents in Library of Congress approved by the 111th, 112th, 113th, or 114th United States Congress. *See* https://www.congress.gov.

[7] According to the data used here, 524 executive agreements were concluded under President Obama during his first term alone.

[8] *See* Hathaway, *supra* note 1.

[9] *Id.* at 1312; *see also* CURTIS A. BRADLEY, INTERNATIONAL LAW IN THE U.S. LEGAL SYSTEM 76 (2015) (pointing out that one of the reasons for the popularity of the executive agreements is that it is "much easier to conclude the growing number of international agreements without submitting them for approval by two-thirds of the Senate").

[10] BRADLEY, *supra* note 9, at 81.

[11] Hathaway, *supra* note 1, at 1285 (arguing that historical conventions explain the use of the treaty). Curtis A. Bradley & Trevor W. Morrison, *Historical Gloss and the Separation of Powers*, 126 HARV. L. REV. 411, 474

https://doi.org/10.1017/ajil.2018.103 Published online by Cambridge University Press

At the same time, anecdotal evidence lends plausibility to alternative explanations as well. Consider, for instance, the bargaining process surrounding arms-reduction agreements between the United States and Russia. During the negotiations of SALT II, the United States proposed a preliminary congressional-executive agreement designed to ban new types of ballistic missiles and cruise missiles. However, former Soviet Foreign Minister Andrei Gromyko rejected the proposal due to the alleged inferior status of the congressional-executive agreement.[12] Similarly, during the negotiations of the Strategic Offensive Reductions Treaty (SORT), the United States and Russia agreed to reduce their arsenal of active nuclear warheads to between 1,700 and 2,200 each. President Putin insisted on codifying the agreement as a formal treaty and spent considerable bargaining power on persuading President Bush, who favored a congressional-executive agreement.[13] Outside the context of nuclear disarmament, negotiation partners have also pointed to the treaty as the desired, more serious form of commitment. For example, when the former President of the Philippines, Corazon Aquino, took office, she voiced her intention to replace the then-current congressional-executive agreements regulating the status of the U.S. military bases in the Philippines by "full-fledged" treaties.[14]

If treaties and congressional-executive agreements are not qualitatively different from one another, it seems hard to rationalize why negotiation partners at times display such great interest in the choice of instrument. Consequently, some scholars appear critical of the supposed lack of the treaties' utility. The arguments come in different forms; some suggest that a president's use of the treaty would signal a particularly high level of commitment,[15] others that the struggle for senatorial approval may cause the government to reveal valuable information truthfully,[16] or that the greater stability of senatorial preferences helps to ensure long-term compliance.[17] What all these accounts have in common is an assumption that treaties, although more politically costly than congressional-executive agreements, confer certain benefits on the parties, in turn justifying their continuing existence as a valuable U.S. policy tool.

As of today, the debate surrounding the ongoing relevance of treaties in a context where congressional-executive agreements are so readily available and widely used remains unsettled. This Article seeks to shed light on the question of whether the treaty is a qualitatively different form of commitment than the congressional-executive agreement. It uses the most comprehensive dataset on U.S. international agreements available—the 7,966 agreements reported in the *Treaties in Force* Series from 1982 to 2012. In contrast to previous analyses,

(2012) (arguing that the use of the treaty can at least partially be explained through selective senatorial attention paid to "major" agreements).

[12] Don Oberdorfer, *Incremental Step: Pact Far Short of Carter's Initial Goal*, WASH. POST (May 11, 1979).

[13] Edward Epstein & Anna Badkhen, *U.S., Russia to Slash Nuclear Arsenals/Bush Wins Concessions—Putin Gets Formal Treaty*, SFGATE (May 14, 2002).

[14] Seth Mydans, *Marcos Flees and Is Taken to Guam; U.S. Recognizes Aquino as President*, N.Y. TIMES (Feb. 26, 1986).

[15] John K. Setear, *The President's Rational Choice of a Treaty's Preratification Pathway: Article II, Congressional-Executive Agreement, or Executive Agreement?*, 31 J. LEGAL STUD. S5 (2002); Lisa L. Martin, *The President and International Commitments: Treaties as Signaling Devices*, 35 PRESIDENTIAL STUD. Q. 440 (2005).

[16] Kenneth A. Schultz, *Domestic Opposition and Signaling in International Crises*, 92 AM. POL. SCI. REV. 829 (1998); John Yoo, *Rational Treaties: Article II, Congressional-Executive Agreements, and International Bargaining*, 97 CORNELL L. REV. 1 (2011).

[17] Setear, *supra* note 15; LISA L. MARTIN, DEMOCRATIC COMMITMENTS: LEGISLATURES AND INTERNATIONAL COOPERATION 64 (2000).

https://doi.org/10.1017/ajil.2018.103 Published online by Cambridge University Press

this Article is the first to directly contrast the consequences of relying on treaties versus executive agreements. Using survival time analysis, the Article demonstrates that, on average, an executive agreement made in 1982 had a 50 percent probability of breaking down by 2012, while a comparable promise made as a treaty broke down with only 15 percent probability.[18] This result holds even after controlling for a number of observable characteristics, such as the composition of the House and the Senate, the subject area of the agreement, and the partner country. The findings also reveal that the difference between the instruments is most pronounced when comparing treaties to *ex ante* congressional-executive agreements.

The results are consistent with the view that promises made in the form of the treaty are qualitatively different from those struck as congressional-executive agreements. Against the backdrop of this empirical finding, it seems premature to call for the abandonment of the treaty, which may still serve important policy functions that cannot similarly be fulfilled by the congressional-executive agreement.

The rest of the Article proceeds as follows: Part II lays out the institutional foundation of the different commitment devices and reviews the theories on how treaties may or may not differ from executive agreements. Part III motivates the empirical inquiry in the context of this theoretical debate, describes the data and methodology used in this study, and presents summary statistics. Part IV presents the results of a formal test of instrument durability, while Part V discusses their implications. A last section concludes.

II. Theory

The United States has two different mechanisms for concluding binding international agreements.[19] The first option is the traditional treaty. Treaties follow the advice and consent procedure set forth in Article II of the Constitution, which requires that, while a treaty is negotiated by the executive, it must still be approved by a two-thirds majority in the Senate in order to be ratified and become binding.[20]

The second option is the executive agreement. Executive agreements can further be categorized into different types. *Congressional*-executive agreements require a simple majority in both the House of Representatives and the Senate.[21] They are used in subject areas in which the executive does not have sole competences. Congressional approval can be obtained after the agreement was negotiated, as was the case with the North American Free Trade Agreement (NAFTA)[22] or the Uruguay Round Agreements of the General Agreement on Tariffs and Trade.[23] However, it is much more common for Congress to provide broad authorization to the president *ex ante* in a statute.[24]

[18] For the reasons causing the State Department to consider an agreement out of force, *see infra* at 14.

[19] Hathaway, *supra* note 1, at 1239 (pointing out that "virtually no other country" has a two-track procedure of making international law like the United States does).

[20] U.S. Const. Art. II, § 2, cl. 2.

[21] Their supposed constitutional basis is the subject of debate and will be detailed momentarily.

[22] North American Free Trade Agreement Implementation Act, Pub. L. No. 103-182, 107 Stat. 2057 (1993).

[23] Uruguay Round Agreements Act, Pub. L. No. 103-465, 108 Stat. 4809 (1994).

[24] Hathaway, *supra* note 1, at 1256 (conducting a search for congressional-executive agreements that have been approved *ex post* and finding only a "small number" of such agreements).

https://doi.org/10.1017/ajil.2018.103 Published online by Cambridge University Press

If the executive has the competence to make policy without referring to Congress, the president may use *sole* executive agreements. Such areas encompass, among others, issues under the president's general executive authority or his role as commander in chief of the armed forces.[25] Sole executive agreements do not require congressional approval, but, like congressional-executive agreements, need to be reported to Congress under the Case Act.[26]

The terminology surrounding the different types of executive agreements has sometimes caused confusion. Political scientists rarely distinguish between different types of executive agreements. When the unmodified term "executive agreement" appears in the political science literature, it commonly refers to the collective of both sole and congressional-executive agreements. In contrast, when international legal scholars use the term "executive agreement," they typically refer to sole executive agreements, whereas the collective of both sole and congressional-executive agreements is not associated with any specific term. In order to preserve flexibility and precision in language, the present Article uses modifiers whenever it refers to a specific type of executive agreement. The unmodified term "executive agreement" is used to refer to the collective of both sole and congressional-executive agreements.

From an international legal viewpoint, it is clear that treaties and executive agreements are perfect substitutes. Indeed, international law does not recognize the term "executive agreement." The term "treaty" is more broadly defined than in the domestic context of the United States. The Vienna Convention on the Law of Treaties (VCLT) states that any written agreement between states governed by international law qualifies as a "treaty" and thus creates a binding legal commitment.[27] Since both U.S. treaties and executive agreements meet this definition, there is no legal difference between either of those commitment devices from the perspective of international law. And while the United States is not a party to the VCLT, the State Department effectively applies the VCLT's definition, thus treating both treaties and executive agreements as equivalent under international law.[28]

Domestically, the issue of legal substitutability has traditionally been more controversial. To be sure, there is little argument that congressional participation can be fully removed by replacing the treaty with the *sole* executive agreement.[29] However, views on the

[25] *Treaties and Other International Agreements: The Role of the United States Senate*, 106 COMM. PRINT 5 (2001) (detailing that presidents have claimed as a basis general executive authority in Article II, Section 1 of the Constitution; his power as commander in chief in Article II, Section 2, Clause 1; his treaty negotiation power in Article II, Section 2, Clause 2; his authority to receive ambassadors in Article II, Section 3; and his duty toward the faithful execution of laws in Article II, Section 3).

[26] 1 U.S.C. 112b(a) (1979). It is important not to "fetishize" this triptych of treaties, congressional-executive agreements, and sole executive agreements. Indeed, most recent scholarship has called attention to its unsuitability in categorizing two recent agreements, namely the Paris Climate Change Agreement and the Iran Nuclear Deal. *See* Jean Galbraith, *From Treaties to International Commitments: The Changing Landscape of Foreign Relations Law*, 84 U. CHI. L. REV. 1675 (2017); Harold H. Koh, *Triptych's End: A Better Framework to Evaluate 21st Century International Lawmaking*, 126 YALE L.J. F. 338 (2017). Since this Article is interested in the *substantive* difference between executive agreements and treaties concluded between 1982 and 2012 and does not discuss novel forms of international agreements, there is little need to move beyond this traditional distinction.

[27] *See* Vienna Convention on the Law of Treaties, Art. 2(1)(a), *opened for signature* May 23, 1969, 1155 UNTS 331 [hereinafter VCLT].

[28] *See* Arthur W. Rovine, *Digest of United States Practice in International Law* 195 (Office of the Legal Adviser, Department of State 1974). For a general overview of the history of U.S. agreements under the VCLT, see Maria Frankowska, *The Vienna Convention on the Law of Treaties Before United States Courts*, 28 VA. J. INT'L L. 281 (1987).

[29] BRADLEY, *supra* note 9, at 90 ("Most scholars . . . believe that the president's authority to enter into sole executive agreements is substantially narrower than the president's authority to enter into Article II treaties."); LOUIS

https://doi.org/10.1017/ajil.2018.103 Published online by Cambridge University Press

interchangeability of treaties and *congressional*-executive agreements are less harmonious. The Constitution does not expressly mention the existence of an instrument that resembles today's congressional-executive agreement, resulting in a debate about how to interpret this silence. To early proponents, it was largely sufficient to show that interchangeability offers flexibility and best describes the practice of U.S. foreign policy to assert that treaties and congressional-executive agreements should act as legal substitutes.[30] Later arguments rested on the idea of the existence of "constitutional moments" that would allow constitutional interpretation to be informed by consistent practice of the president, Congress, and the Supreme Court.[31] Such moments, particularly formed through practice in the 1940s, are alleged to have transformed the meaning of the Treaty Clause, providing a constitutional basis to the congressional-executive agreement.

In contrast, opponents of substitutability highlight the lack of clear textual support. According to a view derived from a strict construction of the Constitution, the Treaty Clause is clear in making senatorial advice and consent the exclusive method for the approval of international agreements.[32] An alternative view derived from a more flexible reading of the Constitution holds that treaties and congressional-executive agreements both have their respective areas of applicability. The argument rests on the idea that the U.S. Constitution has conferred limited powers upon Congress and the president and that executive agreements can only be used within this limited scope. Treaties as the default tool for matters in foreign affairs are not similarly constrained. Thus, if a matter of foreign policy falls outside of the competences that have been conferred upon Congress, the treaty is held to be the exclusive instrument through which legally binding commitments can be made.[33]

HENKIN, FOREIGN AFFAIRS AND THE UNITED STATES CONSTITUTION (1996) (describing the view that the president will seek the Senate's approval only for "prudential reasons" as "unacceptable").

[30] *See* MCCLURE, *supra* note 3, 4, 247 (finding that 1,200 of 2,000 agreements have been concluded as congressional-executive agreements and using this as a basis to advance a basis for legitimizing their use); *see also* Quincy Wright, *The United States and International Agreements*, 38 AJIL 341, 354 n. 62 (1944) (reversing previous views based on "Congressional and executive practice"); Bruce Ackerman & David Golove, *Is NAFTA Constitutional?*, 108 HARV. L. REV. 799, 868 (1995) (demonstrating how McClure's narrative makes consistent practice a necessary and sufficient condition for interchangeability. Also discussing Wright's shift in views.). *See generally* Myers S. McDougal & Asher Lans, *Treaties and Congressional-Executive Agreements or Presidential Agreements: Interchangeable Instruments of National Policy: I*, 54 YALE L.J. 181 (1945); Myers S. McDougal & Asher Lans, *Treaties and Congressional-Executive Agreements or Presidential Agreements: Interchangeable Instruments of National Policy: II*, 54 YALE L.J. 534 (1945) (arguing that a need for flexibility justifies perfect interchangeability of treaties and congressional-executive agreements).

[31] *See* Ackerman & Golove, *supra* note 30, at 861.

[32] Edwin Borchard, *The Proposed Constitutional Amendment on Treaty-Making*, 39 AJIL 537, 538 (1945) (describing the rise of the executive agreement as an "encroachment on the treaty-making power"); Raoul Berger, *The Presidential Monopoly of Foreign Relations*, 71 MICH. L. REV. 1, 48 (1972) (criticizing the idea of "adaptation by usage" as grounds for constitutional interpretation); Laurence H. Tribe, *Taking Text and Structure Seriously: Reflections on Free-Form Method in Constitutional Interpretation*, 108 HARV. L. REV. 1221, 1249 (1995) (criticizing Ackerman's extension on interpretive methods; even though he acknowledges that the Constitution is silent on many questions of separation of powers in foreign affairs, Tribe argues that the Treaty Clause is clear in making Advice and Consent the exclusive method for treaty approval).

[33] John C. Yoo, *Laws as Treaties?: The Constitutionality of Congressional-Executive Agreements*, 99 MICH. L. REV. 757 (2001) (arguing against both "transformationists" who introduce the idea of constitutional moments, as well as "exclusivists" who view treaties as the only means to enact binding international agreements).

While some doctrinal criticism against the widespread use of the congressional-executive agreement in place of the treaty remains,[34] today the predominant view is that treaties and congressional-executive agreements act as legal substitutes under domestic law for the vast majority of agreements.[35] This view is also reflected in Restatement (Third) of the Foreign Relations Law of the United States.[36] There, the American Law Institute states:

> The prevailing view is that the Congressional-Executive agreement can be used as an alternative to the treaty method in every instance. Which procedure should be used is a political judgment, made in the first instance by the president, subject to the possibility that the Senate might refuse to consider a joint resolution of Congress to approve an agreement, insisting that the president submit the agreement as a treaty.[37]

The view that treaties and congressional-executive agreements can be considered legal substitutes naturally raises the question why the United States needs two legal instruments to regulate the same types of international relationships. Indeed, some commentators have asked why the United States should not abandon the treaty in favor of the congressional-executive agreement.[38]

In an attempt to answer this question, scholars have put forward several hypotheses with regard to the contemporary role of the treaty. These can be broadly subsumed under two categories: First, there are hypotheses supporting the idea that treaties have no independent value as a policy instrument. These accounts typically explain the use of the treaty through motivations that are orthogonal to considerations pertaining to the quality of the promise itself. Second, there are hypotheses suggesting that promises made as treaties are qualitatively different from those made as executive agreements, and that the choice to use the treaty is governed by these considerations.

It is important, however, not to fetishize this dichotomy. With few exceptions, neither the hypotheses described nor their respective authors purport to explain the executive's motivations in ways that result in particularly strong policy recommendations. Instead, in formulating the hypotheses, most commentators are open to the idea that the role of the treaty can only be fully explained when taking several of the suggested mechanisms into consideration.

[34] There is some remaining debate as to whether congressional-executive agreements can be used in the rare occasions where the agreement falls outside of the congressional powers enumerated in Article I of the Constitution. *See, e.g.*, Hathaway, *supra* note 1, 1339.

[35] Koh, *supra* note 5, at 91–93 (describing perfect legal substitutability as the "long-dominant view" and pointing out that legal academia rejected opposing conclusions); Koh, *supra* note 26, at 339 (describing the debate as "long ago settled").

[36] The approved draft of the Restatement (Fourth) is conspicuously silent on the matter of interchangeability, but there is no indication that this silence provides support to those arguing against interchangeability. The drafters of Restatement (Fourth) make clear that they focus on Article II treaties only and leave other international agreements unaddressed. *See* Restatement (Fourth) of the Foreign Relations Law of the United States, § 113, Reporters' Note 8 (Mar. 20, 2017). So far, there seems to be little indication of a change in the scholarly debate. *See, e.g.*, Curtis A. Bradley, *Exiting Congressional-Executive Agreements*, 67 DUKE L.J. 1615 (2018) (viewing treaties and congressional-executive agreements as largely interchangeable even after the approval of the draft of Restatement (Fourth)).

[37] Restatement (Third) of the Foreign Relations Law of the United States, § 303, cmt. e (1987).

[38] *See* MCCLURE, *supra* note 3, at 363 (reducing the treaties' relevance to a small subset of non-controversial issues); *see also* LOUIS HENKIN, CONSTITUTIONALISM, DEMOCRACY, AND FOREIGN AFFAIRS 60 (1990) (finding that the executive agreement is the more democratic tool); *see also* Ackerman & Golove, *supra* note 30, at 916 (concluding that the rise of the congressional-executive agreement promotes "[e]fficacy, democracy [and] legitimacy").

https://doi.org/10.1017/ajil.2018.103 Published online by Cambridge University Press

Hypotheses Suggesting Qualitative Equivalence

One of the most influential and forceful accounts against the notion that treaties retain a special policy purpose, and that they may in fact be inferior to the congressional-executive agreement, is provided by Oona Hathaway.

The argument centers on the ease with which a president may renege on an agreement *after* it has been concluded. In particular, Hathaway suggests that the treaty form hinders the ability of presidents to credibly tie their hands, because even after ratification the treaty offers two additional possibilities to renege on a promise that the congressional-executive agreement would not provide. This in turn makes it difficult for other countries to rely on commitments made in the form of the treaty.

The first opportunity to renege is rooted in the fact that non-self-executing treaties follow a two-step process to become enforceable in U.S. law.[39] After ratification, non-self-executing treaties require additional implementation through a legislative act for which a simple majority in both the House and the Senate is required. Compare this to the executive agreement, for which the implementing legislation can and often is adopted in the same step as the ratification. Hathaway argues that the treaties' two-step process makes it possible for the president to renege on his promise after ratification, whether intentionally or because the domestic political costs of enacting the implementing legislation are too high.[40]

The second opportunity is that presidents can withdraw from treaties unilaterally, whereas the withdrawal from congressional-executive agreements requires congressional participation.[41] This would allow presidents to easily renege on their promise even after a treaty has gone through the advice and consent process.

Hathaway supplements her theory with an empirical assessment of 3,119 agreements concluded between 1980 and 2000. She finds that the observed pattern of treaty use is incompatible with theories that ascribe a different quality to the treaty *vis-à-vis* the congressional-executive agreement. Instead, Hathaway argues that the use of the treaty instrument can best be explained through a historical lens.[42] Under her view, the prevalence of the congressional-executive agreement is the result of Congress' desire to reduce trade barriers in the post-World War II era, which necessitated giving the president more flexibility and authority in negotiating trade agreements.[43] This has then led to the conventional use of the congressional-executive agreements in trade (and financial) matters. In other subject areas such as human rights,

[39] Hathaway, *supra* note 1, at 1317.

[40] *Id.* at 1319.

[41] *Id.* at 1336 ("[T]he President is on the whole likely to find it more difficult to withdraw unilaterally from a congressional-executive agreement than an Article II treaty."). Some scholars cast doubt on the claim that presidents can withdraw from treaties more easily than from congressional-executive agreements. As Koremenos and Galbraith point out, many agreements in the UN Treaty Collection have withdrawal provisions that would allow a president to lawfully exit an agreement regardless of the form in which it has been concluded. *See* BARBARA KOREMENOS, THE CONTINENT OF INTERNATIONAL LAW: EXPLAINING AGREEMENT DESIGN 124 (2016) (finding that 70% of agreements have withdrawal provisions, based on a random sample of treaties in the UN Treaty Collection); Galbraith, *supra* note 26, at 1720 (arguing that withdrawal provisions give successors of the current president an easy way to legally withdraw from a treaty). For doctrinal challenges to the claim that treaties can be withdrawn from more easily, see Bradley, *supra* note 36.

[42] Hathaway, *supra* note 1, at 1239–40 ("Although there are patterns to the current practice of using one type of agreement or another, those patterns have no identifiable rational basis.").

[43] *Id.* at 1304.

https://doi.org/10.1017/ajil.2018.103 Published online by Cambridge University Press

the debate was highly politicized and Congress had no desire to give up what was perceived as the nation's sovereignty subject to the lower legislative bar set by the congressional-executive agreement.

To Hathaway, the treaty is a less reliable instrument and should be abandoned in favor of the congressional-executive agreement. Her claim sparked substantial discussion among international law scholars. For example, in 2014, under the title *The End of Treaties?*, the online companion of the *American Journal of International Law* published several essays by prominent scholars and State Department officials discussing whether the treaty will have any place in the future of U.S. foreign policy.[44]

Other scholars have similarly relied on historical perspectives in explaining the use of the treaty. For instance, Curtis Bradley and Trevor Morrison have proposed that the presidents' choice between the two commitment devices is at least partially influenced by a continuous and concerted insistence of the Senate to retain an important role in the ratification process.[45] Such an insistence, mainly expressed through declarations and correspondence, would create a "soft law" that imposes political constraints on the options available to the executive. At the same time, senatorial attention is selective, with a primary focus on "major" agreements in which both the stakes and the public attention are exceptionally high. Hence, differences in the use of the treaty would be at least partially explained by the fact that some agreements, in particular major arms-control agreements, are subject to senatorial scrutiny, whereas other are of less concern to the Senate. The authors provide anecdotal support for their hypothesis, recalling instances in which presidents have abandoned plans to conclude major arms-control agreements such as the Treaty on Armed Conventional Forces in Europe as congressional-executive agreements after pressure by the Senate.

Another account explaining the presidents' choice between the two instruments is the "evasion hypothesis."[46] Particularly prevalent in the writings of political scientists, this reasoning holds that the president's main motivation for choosing one instrument over the other is presidential support for the agreement in the Senate.[47] If an agreement is easy to push through the Senate, the argument goes, presidents will rely on the treaty. If, however, securing a two-thirds majority proves difficult, the president, according to this argument, can simply switch to the congressional-executive agreement without any significant consequences.

Empirical support is provided by Margolis, who analyzes all international agreements concluded from 1943 to 1977 and finds that the distribution of seats in the Senate is highly predictive of the choice between treaties and congressional-executive agreements.[48] The findings support the evasion hypothesis, according to which the choice between treaties and executive agreements is solely a function of domestic legislative support.

[44] 108 AJIL UNBOUND (2014), *available at* https://www.cambridge.org/core/journals/american-journal-of-international-law/ajil-unbound; *see also* BRADLEY, *supra* note 9, at 85 (agreeing with Hathaway that the different use of treaties and executive agreements does not reflect any discernable logic).

[45] Bradley & Morrison, *supra* note 11, at 474.

[46] So labeled by MARTIN, *supra* note 17, at 53.

[47] Terry M. Moe & William G. Howell, *The Presidential Power of Unilateral Action*, 15 J. L. ECON. & ORG. 132, 163 (describing how simply labeling an agreement "executive agreement" rather than "treaty" would allow the president to set foreign policy without having to involve the Senate); *see also* MATTHEW A. CRENSON & BENJAMIN GINSBERG, PRESIDENTIAL POWER: UNCHECKED AND UNBALANCED 321 (2007) (arguing that Franklin D. Roosevelt's use of the executive agreement was motivated by a desire to circumvent the Senate).

[48] LAWRENCE MARGOLIS, EXECUTIVE AGREEMENTS AND PRESIDENTIAL POWER IN FOREIGN POLICY 45 (1986).

https://doi.org/10.1017/ajil.2018.103 Published online by Cambridge University Press

Hypotheses Suggesting Qualitative Difference

In contrast to the hypotheses mentioned in the previous subsection, several accounts suggest that promises made in the form of a treaty are qualitatively different than those made as congressional-executive agreements. These accounts rest on the view that the treaty, although more politically costly, may also confer certain benefits on the parties, which ultimately may lead to a more robust commitment. In interactions where the benefits outweigh the costs, the treaty would then be the preferable instrument, whereas a congressional-executive agreement would be preferred in others.

An account that ascribes political benefits to the treaty is illustrated by the work of John Setear[49] and Lisa Martin.[50] Their reasoning focuses on the high legislative hurdles of the treaties' advice and consent procedure. Since presidents typically lack enough support in the Senate to secure a two-thirds majority, they often have to go through a substantial political struggle to convince senators to vote in favor of a proposed treaty. This political struggle demands not only time and resources, but securing sufficient support may also require the president to make substantial concessions in other subject areas.[51] Because the conclusion of a treaty comes at such a high political cost, Setear and Martin argue that only presidents who are especially committed to the agreement would be willing to go through the advice and consent procedure. If a high level of commitment is not required, the president would instead opt for the sole or congressional-executive agreement,[52] which, as the authors allege, comes at a lower cost. Other countries are aware of this signaling dynamic. When contracting with the United States, they would thus observe the form of agreement that is proposed and, in some high-stakes scenarios, they may refuse to agree unless the president is willing to commit via treaty.

Martin conducts an empirical analysis to support the signaling theory. She provides an analysis of 4,953 international agreements concluded between 1980 and 1999 and finds that the value of the underlying relationship governed by the agreement is determinative of whether a president uses a treaty or an executive agreement.[53] Value is proxied using an indicator for whether the agreement is multilateral, the GNP per capita of the contractual partner, as well as the total GNP.[54] Martin finds that presidents are especially likely to rely on the treaty if the underlying value of the relationship is high. She concludes from these findings that the treaty is reserved for high stakes negotiations in which the president needs to signal a strong commitment to treaty partners.

Advocates of the signaling theory model the bargaining process surrounding the conclusion of international agreements as a signaling game with three rounds: First, it is determined at random whether the president is reliable or not. Then, the president chooses the treaty or the executive agreement. Lastly, the negotiation partner chooses whether to accept or reject the

[49] Setear, *supra* note 15 (describing a signaling model in non-formal terms).

[50] Martin, *supra* note 15 (providing a formal signaling model in which the cost of the agreement determines its credibility).

[51] For a detailed discussion of the supposed difference in political costs, *see* Setear, *supra* note 15, at S14.

[52] Martin believes the cost differential to be greatest for the difference between sole executive agreements and treaties, but she extends her argument to the difference between congressional-executive agreements and treaties. *See* Martin, *supra* note 15, at 447.

[53] *Id.* at 456.

[54] For a detailed description of these proxies, see *id.* at 454.

https://doi.org/10.1017/ajil.2018.103 Published online by Cambridge University Press

proposed agreement and the parties pay their costs and receive their benefits. In reality, international cooperation is more complicated. In particular, signaling theory implicates only the commitment level of the negotiating president, even though many agreements are intended to and do outlast presidential terms. It is unclear from the signaling model why negotiation partners should put much faith into promises made as treaties by one administration when future administrations can easily renege on the agreement.

Addressing some of these limitations, a second hypothesis put forth is that the treaty's high legislative hurdles help solve commitment problems arising out of executive turnover. They purport that the strong legislative support implicit in the treaty mechanism reassures negotiation partners that the United States is likely to cooperate in the long run, even if administrations change.[55] This rationale hinges on the assumption that senatorial preferences are more stable than the preferences of the presidency, for example, because the Senate represents a broader consensus among the voting population that is less sensitive to political shocks,[56] or because senators serve longer terms and avoid changing their position in order to not be seen as wavering.[57] This would in turn allow other countries to rely more heavily on a promise made in the form of a treaty.

Lastly, some scholars have argued that a key difference between treaties and congressional-executive agreements lies in the information that is produced in the process of securing legislative approval.[58] That is, in the course of concluding an agreement as a treaty, the executive needs to reveal important private information in order to make a convincing case and ensure approval of a qualified majority in the Senate. This dynamic can be illustrated by borrowing the leading example posited by John Yoo, who considers a potential military conflict between the United States and China over a territory and negotiations surrounding how this territory would be divided up. The domestic struggle for approval of a treaty requires negotiators to truthfully indicate to the Senate the chances of them winning a war against China. Yoo argues that observing this process would allow China to gain more accurate information about U.S. beliefs than the congressional-executive agreement provides. China may thus insist on the agreement being concluded in the form of a treaty and because the underlying information is more accurate, be incentivized to put more trust into continuing compliance with the agreement.

III. Motivation and Empirical Approach

The foregoing discussion reveals that there are myriad theories on the qualitative difference between treaties and congressional-executive agreements. A common approach to resolving such theoretical debates is to focus on empirical evidence. However, as has been pointed out above, the theoretical and empirical literature both seem inconclusive, with hypotheses of both categories claiming to be supported by quantitative empirical evidence.

[55] Setear, *supra* note 15, at S16; Martin, *supra* note 17, at 64; Eric A. Posner & Jack L. Goldsmith, *International Agreements: A Rational Choice Approach*, 44 Va. J. Int'l L. 113, 124 (2003).

[56] Posner & Goldsmith, *supra* note 55. at 124.

[57] James D. Fearon, *Domestic Political Audiences and the Escalation of International Disputes*, 88 Am. Pol. Sci. Rev. 577 (1994) (detailing audience costs for presidents who back down). Martin, *supra* note 17, at 64 (applying the audience cost rationale to senators).

[58] Schultz, *supra* note 16; Yoo, *supra* note 16, at 25; Posner & Goldsmith, *supra* note 55, at 124.

https://doi.org/10.1017/ajil.2018.103 Published online by Cambridge University Press

It is possible to make sense of this ambiguity in empirical results by observing that previous studies all follow a similar approach. The researcher analyzes the environment in which an international agreement has been concluded and tries to identify patterns that help predict the type of instrument that has been used. If the pattern is consistent with a motivation that assigns different significance to treaties and congressional-executive agreements, that is taken as evidence that these instruments differ in their quality. For instance, from the finding that treaties are used when the partner country has a high GDP per capita, Martin infers that treaties are preferred when the stakes are high and must thus be more reliable commitment devices than congressional-executive agreements. Similarly, from the finding that few treaties are concluded in the area of trade, Hathaway infers that the choice to use the treaty must be animated by historical conventions that made the congressional-executive agreement attractive for trade negotiations.

This focus on choice patterns is a very indirect approach to identifying differences in policy instruments that rests on a number of strong assumptions, such as a correct model specification and a causal relationship between identified patterns and hypothesized motives. Without making these assumptions, observed actions can be the result of a great number of different motivations, making it impossible to infer which instrument is more reliable. For instance, Martin's finding that high GDP correlates with treaty use does not necessarily imply that the treaty is used *because* the partner country has a high GDP. As will be shown below, the treaty is common[59] for agreements between the United States and Western European countries.[60] On average, these countries tend to have a high GDP per capita, but they also share a number of other characteristics that could potentially explain the results, such as a shared history of Roman law and an adherence to legal formalism. In addition, Martin's findings that "high value" agreements increase the probability of treaty use are not only consistent with signaling theory, but are also consistent with the hypothesis proposed by Bradley and Morrison of selective senatorial attention paid to "major" agreements.

Similarly, Hathaway's analysis is purely descriptive and is not equipped to empirically investigate the reasons why treaties are more or less common in certain subject areas. This makes it impossible to determine whether historical conventions motivate the use of the treaty in different subject areas or whether other considerations might be at play.

There are other, possibly more profound difficulties with the descriptive results as well. Hathaway's primary empirical guidepost is a comparison of treaties across subject areas. For instance, she finds that there are twenty-seven commercial treaties and only eight environmental treaties, thus suggesting that treaties are more important in commerce than in the area of environmental protection.[61] However, such a comparison of absolute numbers can be misleading, as it does not take into account the overall number of agreements within a subject area. For instance, while the present study also finds that the absolute number of commercial treaties is higher than that of environmental treaties, there are also more commercial executive agreements than environmental executive agreements. Indeed, twenty out of 216 environmental agreements are concluded as treaties, which is a share of 9 percent. However, only

[59] In relative terms.

[60] *See infra* Table 3, which identifies the proportion of agreements made in the form of a treaty for a number of different countries. See also Table 2 in the online appendix for a full list of countries.

[61] Hathaway, *supra* note 1, at 1258 (presenting a table of absolute treaty usage).

https://doi.org/10.1017/ajil.2018.103 Published online by Cambridge University Press

thirty-five out of 783 commercial agreements are treaties, which is a lower proportion of 4 percent. Thus, treaties may in fact be a more important commitment device for environmental agreements than for commerce, even though the absolute numbers suggest otherwise.

In contrast to previous studies, this Article takes a more direct approach to compare treaties and congressional-executive agreements that does not require equally strong assumptions. At the heart of the inquiry into the differences between the two policy instruments lies one simple question: If a given contract between the United States and a partner country is concluded as a treaty, will this lead to a different outcome than if the agreement is concluded as a congressional-executive agreement? If the answer is yes, then this suggests that the treaty is qualitatively different from the congressional-executive agreement. If the answer is no, then treaties and congressional-executive agreements are substantively similar and their use might be motivated solely by circumstances that are irrelevant to the substantive characteristics of the agreement. It is thus instructive to shift the empirical focus and consider whether the use of the treaty is associated with an outcome that differs from the congressional-executive agreement.

Outcomes of international agreements can be compared on a number of dimensions. One possible measure is the level of compliance with an agreement. However, comparing agreements based on compliance rates has several disadvantages in this context. Not only is "compliance" difficult to define, it is also notoriously hard to measure and verify in most contexts.[62] Even if it were possible to accurately measure compliance, it would still leave open the question of how to compare levels of compliance across different agreements in different subject areas.[63] Motivated by the theoretical work previously discussed, this Article instead compares treaties and congressional-executive agreements based on the strength of the commitment associated them, measured as durability.

Using durability as a proxy for commitment strength is justified for three reasons. First, consider an alternative concept of commitment strength—the ability for an agreement to withstand shocks in the political or economic environment.[64] The probability for shocks to occur increases with time, and therefore agreements which are more resistant to changing circumstances are also those that last longer. Hence, durability is positively correlated even with this alternative concept of commitment strength. Second, from a purely practical perspective, the duration of a treaty can be measured objectively, whereas the competing concept of commitment strength would require making a number of subjective decisions, such as about the severity of the shock and the extent to which the agreement did or did not withstand external pressures.[65] Third, different theories use the concepts of

[62] For a discussion on the importance and difficulty of measuring compliance with international agreements, see George W. Downs, David M. Rocke & Peter N. Barsoom, *Is the Good News About Compliance Good News About Cooperation?*, 50 INT'L ORG. 379 (1996).

[63] For instance, it is difficult to compare a breach of a tax treaty to compliance with a nuclear weapons reduction treaty. *See* Beth A. Simmons, *Treaty Compliance and Violation*, 13 ANN. REV. POL. SCI. 273 (2010) for examples of the fragmented nature of studies on treaty compliance.

[64] George W. Downs & Michael A. Jones, *Reputation, Compliance, and International Law*, 31 J. LEGAL STUD. S95, S104 (2002) (presenting a model built on the notion that reliability is the ability to perform even in light of shocks).

[65] For one attempt at codifying the propensity for shocks to occur by issue area, as well as for a discussion of the downsides of this approach, see Barbara Koremenos, *Contracting Around International Uncertainty*, 99 AM. POL. SCI. REV. 549, 554 (2005).

https://doi.org/10.1017/ajil.2018.103 Published online by Cambridge University Press

commitment strength and durability interchangeably, suggesting that both concepts can be viewed as substitutes.[66]

If the treaty does not retain a special value as a policy tool, it should be the case that a promise concluded as a congressional-executive agreement is just as durable as a promise concluded as a treaty. If, on the other hand, treaties are qualitatively different promises, then the average treaty should outlast the average congressional-executive agreement. In this way, the theoretical debate leads to observable and verifiable empirical claims.

The Data

The dataset consists of all agreements reported in the *Treaties in Force* (TIF) series that were signed and ratified between 1982 and 2012.[67] TIF is the official collection of international agreements in force maintained by the U.S. Department of State. It includes information on the signing date, the parties, and the subject area of the agreement, as well as on when the agreement went into force. The agreements in TIF appear in the *Kavass' Guide of Treaties in Force* (Guide).[68] The Guide is an annual publication accompanying TIF, first published in 1982, that contains additional useful information, such as treaty subject matter, a short description, and the parties to the agreement. TIF uses an elaborate but partially incoherent system to categorize agreements by subject area.[69] In total, there are 197 different subjects in the dataset, many with single-digit observations. This Article reduces the dimension of these subject areas to thirty-eight thematically coherent categories, that are detailed in the online appendix.

Of primary relevance to this analysis is the fact that the Guide contains a list of treaties that were indexed in TIF in the year preceding the year of publication, but are no longer indexed thereafter. The Guide thus makes it possible to determine which agreements have been deleted from the TIF and the year in which the deletion took place. An agreement that was listed in TIF in the previous year but is not listed in the current year is considered by the U.S. State Department to be no longer in force.[70] Deletions are based on one of four

[66] Martin, *supra* note 15, at 448 ("At times, U.S. allies demand that long-standing executive agreements be transformed into formal treaties, explicitly stating that such changes would signal U.S. long-term commitment."); Yoo, *supra* note 16, at 41 ("[T]his reading of the Constitution removes from the nation's tool chest an instrument that could . . . lead to the most durable international agreements."); Hathaway, *supra* note 1, at 1316 ("[T]he bar in Congress is generally higher for Article II treaties—which might be thought to create a stronger assurance of political durability.").

[67] Treaties in Force: A List of Treaties and other International Agreements of the United States (1929–2017) [hereinafter TIF]. The dataset is limited to agreements concluded since 1982 because its construction requires cross-references with the *Kavass' Guide of Treaties in Force*, which was first published in 1982. For the first time since 1957, the State Department did not release a separate edition of TIF for the years 2013 and 2014, which makes it impossible to tell the exact year in which agreements went out of force during that two-year window. The analysis thus ends in 2012.

[68] *See* U.S. Department of State, A Guide to the United States Treaties in Force (Igor I. Kavass ed., 1982–2016).

[69] *See*, e.g., U.S. Department of State, A Guide to the United States Treaties in Force, at vii (Igor I. Kavass ed., 2016) ("[T]here is very little correlation between the bilateral subject categories and the multilateral subject headings. The *Treaties in Force* does not have either a numerical or a subject list of bilateral and multilateral agreements in force. Neither does it attempt to draw agreements together in other manners of retrieval convenient to researchers.").

[70] This procedure is accurate, save for some exceptions caused by idiosyncrasies in the publication process. For instance, the Guide (2011) lists the START I agreement as having been indexed in TIF (2010) and not indexed in TIF (2011), even though the agreement expired on December 5, 2009 (The corresponding identifier is KAV 3172, see U.S. Department of State, A Guide to the United States Treaties in Force 870 (Igor I. Kavass

https://doi.org/10.1017/ajil.2018.103 Published online by Cambridge University Press

grounds, though TIF does not indicate the reason for any particular deletion. These grounds are (1) expiration based on the terms of the agreement; (2) denunciation; (3) replacement by another agreement; or (4) termination.[71]

Although TIF lumps these four grounds together, some observers might contend that the expiration of an agreement based on its original terms should be treated differently from the other grounds for deletion. After all, an agreement that goes out of force based on an expiration clause does not appear to "break down" in a way that is comparable to an agreement ending, for instance, through withdrawal. However, whether the parties include a termination provision specifying an expiration date in an agreement is endogenous, meaning that it may be at least partially determined by the same circumstances that lead to traditional breakdown of an agreement. In the related area of private contracts, studies have shown that a contract with an unreliable negotiation partner not only increases the probability that an agreement will be terminated prematurely, but also increases the probability that a termination clause with a limited duration is included in the contract.[72] Thus, both agreements that go out of force due to termination, withdrawal or replacement, and those that end according to expiration clauses, implicate a party's reliability.

For each agreement, the Guide further reports a "Senate Treaty Document Number." This number is assigned to any treaty submitted to the Senate under the advice and consent procedure. Executive agreements do not receive a Senate Treaty Document Number. The number can thus be used to identify which agreement in the database is a treaty and which agreement was concluded as an executive agreement.

In order to distinguish between agreements that Congress authorized *ex post* and other types of executive agreements, the data further relies on Hathaway's collection of authorizing legislation. Hathaway compiled the most comprehensive list of congressional acts between 1980 and 2000 that can reasonably be construed as authorizing prior congressional-executive agreements. This list encompasses nine legislative acts.[73] I inspect each of these acts manually and then search the TIF data for executive agreements that can reasonably be construed as being authorized under one of the acts.[74] This approach identifies fifty-two *ex post*

ed., 2011)). This is due to the fact that the treaty expired too close to the TIF's 2010 publication deadline. However, all agreements are equally affected by idiosyncrasies in the underlying publication mechanism, which makes it unlikely for these errors to introduce biases in the estimation.

[71] *See, e.g.*, Treaties in Force: A List of Treaties and other International Agreements of the United States, at i (1982).

[72] For empirical evidence, see Keith J. Crocker & Scott E. Masten, *Mitigating Contractual Hazards: Unilateral Options and Contract Length*, 19 RAND J. ECON. 327 (1988). For a theoretical assessment with regards to treaty-making, see Laurence R. Helfer, *Exiting Treaties*, 91 VA. L. REV. 1579 (2005) (describing that an easy termination of a treaty may make its conclusion more attractive for those who do not intend to comply regardless of cost).

[73] S*ee* Hathaway, *supra* note 1, at note 49. It is possible that she misses some acts, as the process requires a manual search of the Statute at Large; however, it is the most comprehensive list to date.

[74] For each statute, a manual search is conducted for all executive agreements between the United States and the party mentioned in the act within a two-year window prior to the passing of the legislation. Not every act actually authorizes an executive agreement in the treaty. For example, the South African Democratic Transition Support Act of 1993 encourages investments and trade in South Africa. *See* South African Democratic Transition Support Act of 1993, Pub. L. No. 103-149, 107 Stat. 1503 (1993). It can thus reasonably be construed as authorizing prior investment and trade agreements between the United States and South Africa. However, there has not been any such agreement in the years preceding the act that is included in TIF. Indeed, the first investment agreement was concluded shortly after the statute was enacted.

https://doi.org/10.1017/ajil.2018.103 Published online by Cambridge University Press

congressional-executive agreement, out of 5,443 agreements that have been concluded between 1982 and 2000.

The dataset was further complemented with publicly available information on the president who signed an agreement,[75] Senate compositions by party, and "legislative potential for policy change" (LPPC) scores for the Senate, as used by Martin.[76] LPPC scores reflect how difficult it is for a president to push legislation through. A higher LPPC score indicates lower political costs to implement legislation in the Senate. The LPPC score is constructed according to the following formula:

$$LPPC = Seats_{Pres}{}^{*} Unity_{Pres} - Seats_{Opp}{}^{*} Unity_{Opp}$$

Unity refers to voting unity scores published by *Congressional Quarterly*.[77] Higher *Unity* scores indicate more uniform voting patterns. Hence a president with a united majority in the Senate will receive a higher LPPC score than a president with an ideologically fractured majority. Similarly, the more seats a president has, the higher the LPPC score will be.

Overall, the dataset contains 7,966 agreements. In longitudinal form, each agreement is observed once per year while it is in force and once when it goes out of force, leading to a total of 129,518 per-year-per-agreement observations.

Before proceeding to the analysis, it is important to address possible limitations of this dataset. While TIF is the most comprehensive collection of international agreements to date, there is no dataset listing every international agreement the United States has previously concluded.[78] Researchers could attempt to complement TIF with other treaty collections with the aim of creating a more comprehensive list of agreements. However, this is neither advisable nor practical for several reasons.

First, and most important, the only known bias in TIF is the omission of secret agreements, which, if publicized, could threaten national security.[79] However, since these secret agreements are by definition not publicly known, it is likely that they are also missing in other databases. Second, the agreements in TIF all follow a comprehensible selection process:

[75] This choice rests on the rationale that the sitting president at the time when the agreement was signed has the greatest influence on its content. However, all results are substantively identical if instead using a categorical variable for the president under which the agreement went into force. The relevant regressions are included in the online appendix.

[76] Martin, *supra* note 15, at 454 (describing LPPC scores).

[77] *Congressional Quarterly Weekly Report*, CONG. Q. (1982–2012). Unity scores are the share of party members voting with their party in party unity votes. Party unity votes are those votes in which a majority of democrats vote in one direction and a majority of republicans vote in the other direction. For instance, if 85% of democrats voted in favor of a bill and 90% of republicans voted against it, then the voting unity scores for that particular bill would be 0.85 for democrats and 0.9 for republicans. These values are averaged across all roll calls for a given year. Note that unity scores are sometimes reported in percentages, e.g., 85 instead of 0.85 and 90 instead of 0.9. To increase readability of the relevant coefficients, I use unity scores as decimals.

[78] For an overview of possible sources, see Marci Hoffman, *United States*, *in* SOURCES OF STATE PRACTICE IN INTERNATIONAL LAW 529 (Ralph F. Gaebler & Alison A. Shea eds., 2d ed. 2014).

[79] The Case Act provides that these agreements only need to be transmitted "to the Committee on Foreign Relations of the Senate and the Committee on International Relations of the House of Representatives under an appropriate injunction of secrecy to be removed only upon due notice from the President." 1 U.S.C. 112b(a) (1979).

https://doi.org/10.1017/ajil.2018.103 Published online by Cambridge University Press

they are agreements submitted to Congress pursuant to the Case Act and are considered to be in force by the State Department. Combining these agreements with other databases introduces the possibility for unknown selection biases, threatening the interpretability of any findings. Third, TIF uses its own index system, such that agreements in TIF cannot easily be compared to those from other sources. And fourth, previous attempts to combine datasets have resulted in substantially fewer agreements than contained in the dataset used here.[80] For these reasons, this Article suggests that a single dataset based on TIF is preferable to a combination of different sources.

A second limitation of the data is that it is not able to identify individual sole executive agreements. This is important because the question of whether the executive agreement should fully replace the treaty is only raised with regard to congressional-executive agreements; it is generally acknowledged that sole executive agreements are very different policy instruments that fall within the president's power and do not require legislative participation.[81] The reason for the inability to identify sole executive agreements individually is that neither TIF nor the executive agreements themselves indicate their authorizing legislation. Thus, identifying sole executive agreements would require a search for authorizing legislation for each individual executive agreement in the Statute at Large, a process that is infeasibly labor intensive and cannot be easily automated.[82] However, prior studies have found that the proportion of sole executive agreements is minimal, with an estimated share of 5 percent to 6 percent of all international agreements.[83] The analysis described below addresses this limitation through a conservative estimation procedure.[84]

Lastly, the dataset does not include information about the party that is primarily responsible for the agreement's end. All theories summarized above focus on the reliability of the United States as a negotiation partner. Agreements terminated due to reasons unrelated to the U.S. level of commitment should thus not factor into the analysis.[85] At the same time, the identity of the party responsible for treaty termination is unobservable unless the researcher analyzes each termination individually. Even then, identifying the responsible party is often a subjective

[80] Hathaway combines multiple sources, leading to a total number of 3,119 agreements in the period of 1980–2000. *See* Hathaway, *supra* note 1, at 1258–60. In contrast, the dataset used here contains 6,148 agreements in the same period.

[81] *See* note 28.

[82] Hathaway, *supra* note 1, at 1259 ("[S]eparating executive agreements that are congressionally authorized from those that are not requires a painstaking search for authorizing legislation. To determine whether an agreement is a congressional-executive agreement, it is necessary to search the Statutes at Large prior to the date the agreement went into effect for terms related to that subject area. Then it is necessary to read each statute to determine whether it actually authorizes the relevant international agreements.") (footnote omitted).

[83] *See* C.H. McLaughlin, *The Scope of the Treaty Power in the United States II*, 43 MINN. L. REV. 651, 721 (1958) (calculating that 5.9% of agreements between 1883 and 1957 were concluded as sole executive agreements, or "Presidential agreements"); *see also* International Agreements: An Analysis of Executive Regulations and Practices, at 22, Senate Committee on Foreign Relations, 95th Cong., 1st Sess. (1977) (calculating that 5.5% of agreements from 1946–1972 relied exclusively on executive authority).

[84] *Infra* at 84.

[85] For example, commentators have observed recent efforts by developing countries to amend, supersede, or exit from bilateral investment treaties (BITs) with the United States. These efforts are fueled by information about the negative domestic effects of BITs and treaty interpretations that tend to favor investors. For a thorough discussion, see Federico M. Lavopa, Lucas E. Barreiros & Victoria M. Bruno, *How to Kill a BIT and Not Die Trying: Legal and Political Challenges of Denouncing or Renegotiating Bilateral Investment Treaties*, 16 J. INT'L ECON. L. 869 (2013).

https://doi.org/10.1017/ajil.2018.103 Published online by Cambridge University Press

assessment. However, the inability to observe responsibility for agreement breakdown is unlikely to introduce significant biases into any estimates derived from the data. Bias is introduced only if an unobserved variable is related both to the variable of interest and the outcome variable. Here, this implies that bias is introduced only if the probability for the other party to breach the agreement differs between treaties and congressional-executive agreements.[86] However, as pointed out above, only the U.S. reliability is implicated by the distinction between treaties and executive agreements. It would thus be surprising if a partner country's propensity to withdraw was related to the choice of the policy instrument.[87]

Methodology

With each observation in the dataset being an agreement-year, the analysis considers variations in the durability of different types of agreements, holding other characteristics constant. Differences in durability are estimated using survival time analysis. These methods are also referred to as event history studies.[88] Before proceeding, it is helpful to define a few key terms. Survival time analysis is primarily used in the medical sciences using terminology encountered most often in clinical trials. A "subject" is a unit of observation, here an agreement. An "event," "death," or "failure" are synonyms for the occurrence of the incident of interest, here the going-out-of-force of an agreement. The "survival time" is the time period between the start of the observation and the occurrence of the incident, here the period in which an agreement is in force. Agreements that are in force in the last period of observation are considered "right-censored," i.e., with a survival time that has a known lower bound and an unknown upper bound, as one cannot observe how long the agreement ultimately lasts.[89] Finally, a "hazard rate" refers to the probability for an event to occur.

Survival time analysis offers different models to estimate the longevity of an observed subject, each of which has advantages and disadvantages. Model choice is primarily governed by whether the survival times of the analyzed subjects are discrete, such that they can be counted, or continuous, as well as how they are observed.

International agreements can go out of force at any point in time and survival times are thus continuous in nature. However, as described above, survival times are measured only once per year based on publication in TIF. In principal, continuous grouped data allows for the application of both parametric and semi-parametric models. However, as explained in the appendix, the semi-parametric Cox proportional hazard model[90] is the most appropriate for the present scenario,[91] as it is a semi-parametric model that only

[86] Otherwise, the omission simply increases the standard errors of the estimate.

[87] A similar argument applies to agreements falling into desuetude. Whether an agreement is still actively relied on cannot be observed, is a subjective determination and we lack a clear theory for why the rate of active reliance to inactivity should differ for treaties and executive agreements.

[88] Janet M. Box-Steffensmeier & Bradford S. Jones, Event History Modeling: A Guide for Social Scientists 2 (2004) (describing the different terminology that survival models are referred to).

[89] The agreement is in force at least until 2012, possibly longer.

[90] David R. Cox, *Regression Models and Life Tables*, 34 J. Royal Stat. Soc'y 187 (1972).

[91] *See, e.g.*, Lu Tian, David Zucker & L. J. Wei, *On the Cox Model with Time-Varying Regression Coefficients*, 100 J. Am. Stat. Ass'n 172, 172 (2005) ("The most popular semiparametric regression model for analyzing survival data is the proportional hazards (PH) model.") (citation omitted). For examples in international law, see Zachary Elkins, Andrew T. Guzman & Beth A. Simmons, *Competing for Capital: The Diffusion of Bilateral Investment Treaties, 1960–2000*, 60 Int'l Org. 811, 828 (2006) (estimating adoption times for bilateral investment treaties

https://doi.org/10.1017/ajil.2018.103 Published online by Cambridge University Press

relies on few assumptions.[92] The complementary log-log model serves as a robustness check.

Summary Statistics

Table 1 reports summary statistics. As can be seen, 5 percent of all agreements between 1982 and 2012 were concluded in the form of a treaty, making the use of the treaty exceptional. 20 percent of all agreements went out of force at some point during the period of observation. The average agreement was observed to be in force for 15.26 years. Among the agreements that are no longer in force, the average durability is 7.3 years. LPPC scores range from −17 to 17 with a mean of −0.10. On average, 50 percent of the seats in the Senate were held by the president's party at the time the agreement was signed. For 71 percent of agreements, the government was divided, with the White House being held by one party and either the Senate, the House, or both being held by the other. Together, these numbers indicate that the average agreement could not have been adopted in the form of a treaty absent bipartisan support, making the treaty a potentially costly instrument.

Figures 1 and 2 depict histograms indicating the number of executive agreements and treaties split by year and by the signing president. The figures show that the total number of agreements peaked in 1985 and declined thereafter. The relative share of treaties among all agreements was greatest in 2010, with 28 percent of agreements being concluded in the form of a treaty. However, most of these treaties were signed prior to the Obama administration. Indeed, President Obama concluded fewer agreements as treaties than any other president during the period of observation, a finding that has previously been observed by other scholars.[93] Meanwhile, agreements signed under President Clinton include the highest share of treaties, with 7.6 percent. Together, this implies that the use of the treaty varies with the president, though executive agreements are by far the more prevalent instrument across all administrations.

Table 2 depicts a list of selected subject areas and the prevalence of treaties and executive agreements in each area. A full list of agreements by subject is included in the online appendix. The only subject in which treaties are more prevalent than executive agreements is extradition, where 94 percent of agreements are concluded as treaties. A likely explanation for this

using a Cox model); Beth A. Simmons, *International Law and State Behavior: Commitment and Compliance in International Monetary Affairs*, 94 AM. POL. SCI. REV. 819, 823 (2000) (relying on the Cox model to estimate time until states accept commitments under IMF Articles of Agreement Article VIII).

[92] The Cox model is of the form

$$h_i(t|x_i) = h_0(t)e^{x_i'\beta}$$

where i is the individual agreement, t is a period in time, x denotes a set of covariates, and h denotes the hazard rate, i.e. the probability for an event to occur.

[93] Jeffrey S. Peake, *Executive Agreements as a Foreign Policy Tool During the Bush and Obama Administrations*, at 2 (Apr. 16, 2015), *available at* https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2594414. *See also* Curtis A. Bradley & Jack L. Goldsmith, *Presidential Control Over International Law*, 131 HARV. L. REV. 1201, 1211 (2018) ("Both the average number of treaties transmitted per presidential year during [President Obama's] Administration (4.75) and the percentage of treaties receiving Senate consent (39%) are by far the smallest in the modern period measured since President Truman. . . .").

https://doi.org/10.1017/ajil.2018.103 Published online by Cambridge University Press

TABLE 1.
SUMMARY STATISTICS

	Mean	SD	Min	Max	Median	IQR
In Force Year	1996	8.59	1982	2012	1995	15
Treaty	0.05	0.22	0	1	0	0
Event	0.20	0.40	0	1	0	0
Out of Force Year*	1998	7.63	1983	2012	2000	11
Durability	15.26	9.03	0	32	15	16
Durability*	7.30	5.68	0	30	6	8
LPPC	−0.10	9.49	−17	17	0	17
Share Senate	0.50	0.04	0	1	0.50	0
Divided Government	0.71	0.45	0	1	1	1
Multilateral	0.06	0.24	0	1	0	0
Intl Organization	0.01	0.12	0	1	0	0

Summary statistics for the variables used in this dataset. An asterisk indicates that the statistics only include treaties that have gone out of force in the period of observation.

phenomenon is uncertainty concerning the constitutionality of using executive agreements to surrender individuals to foreign nations. The uncertainty stems from *Valentine v. United States*,[94] where the Supreme Court held that extraditions need to be authorized "by act of Congress or by the terms of a treaty." However, *Valentine* did not consider whether extraditions pursuant to congressional-executive agreements are constitutional. Instead, the court considered whether the United States can extradite citizens in the absence of any agreement or legislative authority. As such, the reference to "acts of Congress" may have been "pure dicta."[95] Today, commentators are split on the question of whether extraditions can be authorized by congressional-executive agreement, with some emphasizing a lack of congressional authority over extradition[96] and others interpreting *Valentine* as an explicit acceptance by the Supreme Court of such authority.[97]

Other areas in which treaties are very prevalent include "judicial assistance," agreements to prosecute cross-border crime such as drug trafficking, money laundering, and stolen passports; "taxation," which primarily encompasses double taxation and taxation information sharing agreements; and "property," including agreements on the return of stolen vehicles and the transfer of real estate. Considering only subject areas, it seems difficult to explain the use of the treaty along one coherent dimension. For instance, if we think that treaties are especially prevalent among important agreements, we might expect them to be used

[94] Valentine v. United States ex rel. Neidecker, 299 U.S. 5, 8 (1936).

[95] Ntakirutimana v. Reno, 184 F.3d 419, 437 (5th Cir. 1999) (DeMoss, J., dissenting).

[96] *See* Yoo, *supra* note 33, at 812 (arguing that extradition does not clearly fall under one of the enumerated powers conferred to Congress); *see also* Hathaway, *supra* note 1, at 1346–48.

[97] *See* Alexandropoulos Panayiota, *Enforceability of Executive-Congressional Agreements in Lieu of an Article II Treaty for Purposes of Extradition:* Elizaphan Ntakirutimana v. Janet Reno, 45 VILL. L. REV. 107, 113–14 (2000) (arguing that the Supreme Court in *Valentine* has clearly determined the legality of an extradition pursuant to an executive agreement); *see also* Louis Klarevas, *The Surrender of Alleged War Criminals to International Tribunals: Examining the Constitutionality of Extradition via Congressional-Executive Agreement*, 8 UCLA J. INT'L L. & FOR. AFF. 77, 107 (2003) (providing further cases to support the interpretation that *Valentine* authorizes extradition based on an executive agreement).

https://doi.org/10.1017/ajil.2018.103 Published online by Cambridge University Press

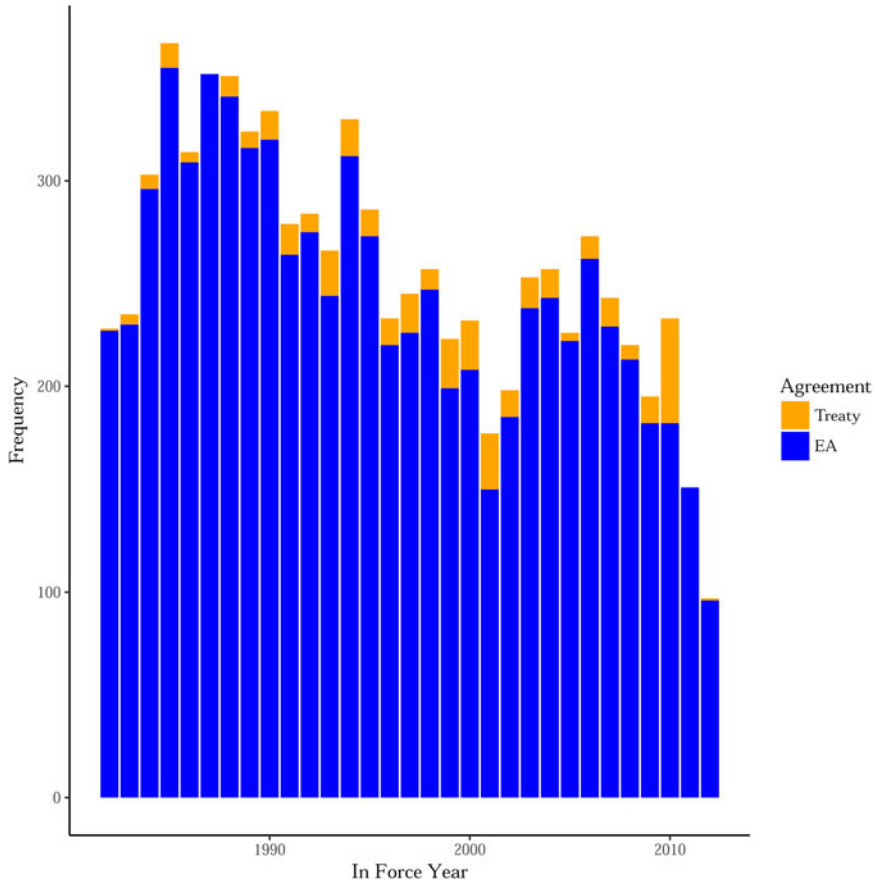

FIGURE 1. Agreement Types Over Time

This figure depicts the use of executive agreements and treaties over time.

frequently in agreements relating to national security and defense.[98] However, only 1 percent of defense agreements are concluded in the form of a treaty. Meanwhile crime prevention, which is often thought of as having a lower priority than national security, includes a much larger share of treaties.

The data also suggest that theories that explain treaty use by reference to historical convention leave many subject areas unexplained. For instance, although some scholars have argued that path-dependence explains why treaties are especially common in human rights and absent in trade, Table 2 shows that neither subject presents a particularly striking outlier. While in the area of human rights, treaties are somewhat prevalent (17 percent of all agreements), the choice of that instrument is still the exception rather than the norm. Similarly, the use of treaties in economic areas such as trade, commerce, and finance is close to the average of 5 percent, raising questions as to whether the rarity of the instrument in these areas is best explained by historical events or whether it instead reflects a different aversion to the treaty

[98] It cannot be ruled out that the importance of agreements within a subject category varies more strongly than the importance of agreements between subject categories. This possibility is discussed below, *infra* at 87–88.

https://doi.org/10.1017/ajil.2018.103 Published online by Cambridge University Press

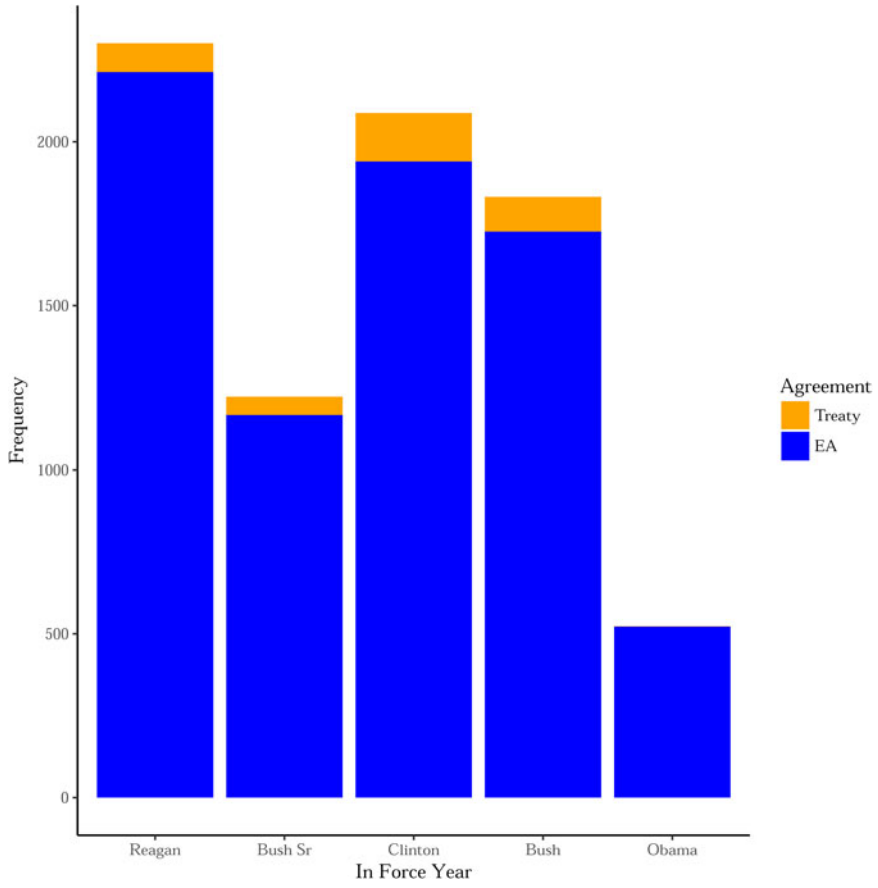

FIGURE 2. Agreement Types by President

This figure depicts the use of executive agreements and treaties by the different presidents.

that affects other subject areas as well. In sum, it seems difficult for conventional theories to explain the wide variety of treaty prevalence in the different subject areas.

Turning to the identity of the parties, the agreements in the dataset were concluded between the United States and one or more of 215 countries or state-like entities and fifty-two international organizations. Table 3 depicts the twenty countries with the most agreements in the dataset. A full list of agreements by partner country is included in the online appendix. The three most frequent treaty partners are all Western European countries, namely France, Italy, and Germany. In multilateral agreements, 20 percent are concluded in the form of a treaty, far exceeding the share in any bilateral relationship.

IV. RESULTS

This section presents the results of the analysis. I first consider differences between treaties and executive agreements in general, without distinguishing among different types of executive agreements. I then investigate whether the findings change when further differentiating

https://doi.org/10.1017/ajil.2018.103 Published online by Cambridge University Press

TABLE 2.
AGREEMENT USE BY SUBJECT AREA

Subject	# EAs	# Treaties	Mean Treaty
Agriculture	454	1	0.00
Education	64	0	0.00
Postal Matters	239	0	0.00
Defense	1433	9	0.01
Other	138	2	0.01
Labor	131	3	0.02
Finance	500	22	0.04
Trade and Commerce	748	35	0.04
U.S. Boundaries	52	4	0.07
IP	23	2	0.08
Environment	196	20	0.09
Fisheries	83	9	0.10
Human and Fundamental Rights	15	3	0.17
Property	8	5	0.38
Taxation	103	75	0.42
Judicial Assistance	93	80	0.46
Extradition	5	75	0.94

The table depicts the prevalence of treaties and executive agreements for selected subject areas. Statistics for all subjects are included in the online appendix.

between sole executive agreements, *ex ante* congressional-executive agreements, and *ex post* congressional-executive agreements.

Treaties vs. Executive Agreements

Table 4 presents results for the Cox proportional hazard model. Model (1) only includes the treaty indicator. It can be thought of as a simple descriptive comparison of the durability of treaties and all executive agreements, taking no other characteristics into account. Model (2) includes president and subject area fixed effects. They indicate the average difference in durability, given that two agreements have been concluded by the same president and in the same subject area. Model (3) additionally includes country fixed effects.[99] If the choice between executive agreements and treaties was the result of historical path-dependence without substantive relevance in the present, then there should be no difference in the durability of treaties and executive agreements when holding all of these characteristics constant. The coefficient on *Treaty* should thus be small and statistically insignificant. Model (4) further controls for the president's party's share of seats in the Senate, as well as for a divided government, in order to investigate whether differences between treaties and executive agreements are explained by a president's intention to side-step the Senate. If that is the primary motivation for choosing the treaty, then the inclusion of either of these covariates should render the coefficient on *Treaty* small and insignificant. Model (5) does not control for the share of seats, but

[99] Due to data sparsity, not all country fixed effects can be accurately estimated, which is why this specification is included separately.

https://doi.org/10.1017/ajil.2018.103 Published online by Cambridge University Press

TABLE 3.
AGREEMENT USE BY PARTNER COUNTRY

Country	# EAs	# Treaties	Mean Treaty
Mexico	247	6	0.02
Japan	250	2	0.01
Russia	219	4	0.02
United Kingdom	195	10	0.05
Canada	190	10	0.05
Egypt	188	2	0.01
South Korea	139	2	0.01
Germany	116	7	0.06
Philippines	116	2	0.02
France	106	10	0.09
Australia	102	4	0.04
China, Republic	104	1	0.01
Indonesia	100	2	0.02
Israel	97	3	0.03
Brazil	98	1	0.01
Ukraine	92	4	0.04
Pakistan	95	0	0.00
Peru	92	1	0.01
Italy	82	6	0.07
Jordan	85	2	0.02

The table depicts the prevalence of treaties and executive agreements for the twenty most frequent partner countries in the dataset. Statistics for all countries are included in the online appendix.

for LPPC scores, which are arguably a better proxy for the costs of pushing legislation through the Senate. The standard errors for all models are clustered by agreement.

In each model specification, the coefficient on the treaty indicator is negative and significantly different from zero. Note that coefficients in survival models express changes in the probability for an event to occur, i.e., the hazard rate. Here, the event is defined as an agreement going out of force. Hence a negative coefficient indicates a decrease in the probability for an agreement to go out of force if it is concluded in the form of a treaty. The results imply that treaties last significantly longer than executive agreements and that the difference in durability is neither the result of arbitrary subject-matter conventions, nor a byproduct of a decision-making process that is primarily driven by the seat map in the Senate.

Table 5 runs the same model specifications using the competing complementary log-log model. Again, the results consistently show that agreements concluded as treaties outlast those concluded as executive agreements. Thus, the results do not depend on the specific idiosyncrasies of the Cox model, but are robust to other model specifications as well.

Statistical significance does not imply substantive relevance. With a large number of observations such as in this study it is important to complement the statistical findings with evidence that the results are substantively meaningful and not only marginal. Differences in survival times can be expressed as hazard ratios, which describe the ratio of the hazard rate for different subgroups. Here, the hazard ratio for the treaty indicator of the preferred

https://doi.org/10.1017/ajil.2018.103 Published online by Cambridge University Press

TABLE 4.
COX PROPORTIONAL HAZARD MODEL

	Dependent Variable: Hazard Rate				
	(1)	(2)	(3)	(4)	(5)
Treaty	−1.324***	−1.313***	−1.164***	−1.176***	−1.180***
	(0.237)	(0.254)	(0.262)	(0.262)	(0.262)
Divided				−0.099	−0.086
				(0.135)	(0.129)
Senate Share				1.423	
				(0.949)	
LPPC					0.008
					(0.004)
President FEs		✓	✓	✓	✓
Subject FEs		✓	✓	✓	✓
Country FEs			✓	✓	✓
Observations	129,518	129,518	129,518	129,518	129,518
Log Likelihood	−12,790	−12,143	−11,905	−11,901	−11,900
Wald Test	31***	23,684***	149,155***	146,607***	146,560***

Note: *p < 0.05; **p < 0.01; ***p < 0.001

The results of a Cox proportional hazard regression of survival time on a treaty indicator and several covariates. Standard errors are clustered by agreement.

Model (5) is 0.3, indicating the relative probability for a treaty to go out of force at any point during the window of observation is about 30 percent of the probability for an executive agreement to go out of force.

The findings can further be illustrated by comparing estimated survival curves or cumulative hazard curves. The survival curve at time *t* depicts the probability that an agreement lasts through *t*, conditional on having lasted up until *t*. The cumulative hazard in time *t* is the probability that an agreement goes out of force in or prior to *t*.

Figure 3 depicts estimated survival and cumulative hazard curves for the preferred Model (5), one corresponding to a treaty and one corresponding to an executive agreement. Numerical covariates have been centered around their mean. For categorical variables, the most prevalent value is used.[100]

Figure 3 reveals that there is a probability of 14 percent for an agreement to break down at the end of the period of observation, conditional on having remained in force until then. For executive agreements, that probability is 40 percent. Similarly, there is a 15 percent probability that a treaty breaks down at some point between 1982 and 2012, whereas that probability is 50 percent for executive agreements.

[100] The country is Mexico, the president is Reagan, and the subject is defense.

https://doi.org/10.1017/ajil.2018.103 Published online by Cambridge University Press

TABLE 5.
COMPLEMENTARY LOG-LOG MODEL

	Dependent Variable: Hazard Rate				
	(1)	(2)	(3)	(4)	(5)
Treaty	−1.324***	−1.314***	−1.164***	−1.176***	−1.180***
	(0.238)	(0.270)	(0.283)	(0.283)	(0.283)
Divided				−0.099	−0.086
				(0.128)	(0.123)
Senate Share				1.426	
				(0.939)	
LPPC					0.008
					(0.004)
Constant	−20.473	−22.219	−37.915*	−38.682	−37.993
	(26.011)	(18.997)	(25.876)	(19.619)	(32.866)
President FEs		✓	✓	✓	✓
Subject FEs		✓	✓	✓	✓
Country FEs			✓	✓	✓
Interval FEs	✓	✓	✓	✓	✓
Observations	129,518	129,518	129,518	129,518	129,518
Log Likelihood	−7,708	−7,061	−6,822	−6,818	−6,818
Akaike Inf. Crit.	15,484	14,269	14,224	14,220	14,219

Note: *p < 0.05; **p < 0.01; ***p < 0.001

The results of a generalized linear model with a complementary log-log link function regressing survival time on a treaty indicator and several covariates. Standard errors are clustered by agreement.

Different Types of Executive Agreements

Thus far, the analysis has not distinguished between different types of executive agreements. There are, however, important differences among these instruments, although these differences have not adequately been taken into account in prior empirical scholarship on this topic. *Ex post* congressional-executive agreements require congressional approval of the individual agreement. Martin's theory implies that this requirement of individual approval may decrease the difference in cost as compared to a treaty. In addition, sole executive agreements are very different policy instruments that fall entirely into the president's power and do not require legislative participation. As such, it may be reasonable to argue that sole executive agreements should be omitted from the analysis. Both of these issues will be addressed in turn.

Consider first *ex post* congressional-executive agreements. As pointed out above, *ex post* congressional-executive agreements are rare, with a share of less than 1 percent during 1980 and 2000. Table 6 presents results on the coefficient of the treaty indicator when the model is run separately on *ex post* congressional-executive agreements and all other international commitments.[101] While most model specifications indicate that there may be a

[101] Note that the analysis only considers agreements concluded between 1982 and 2000 to take into account the fact that *ex post* congressional-executive agreements are not identified beyond that window.

https://doi.org/10.1017/ajil.2018.103 Published online by Cambridge University Press

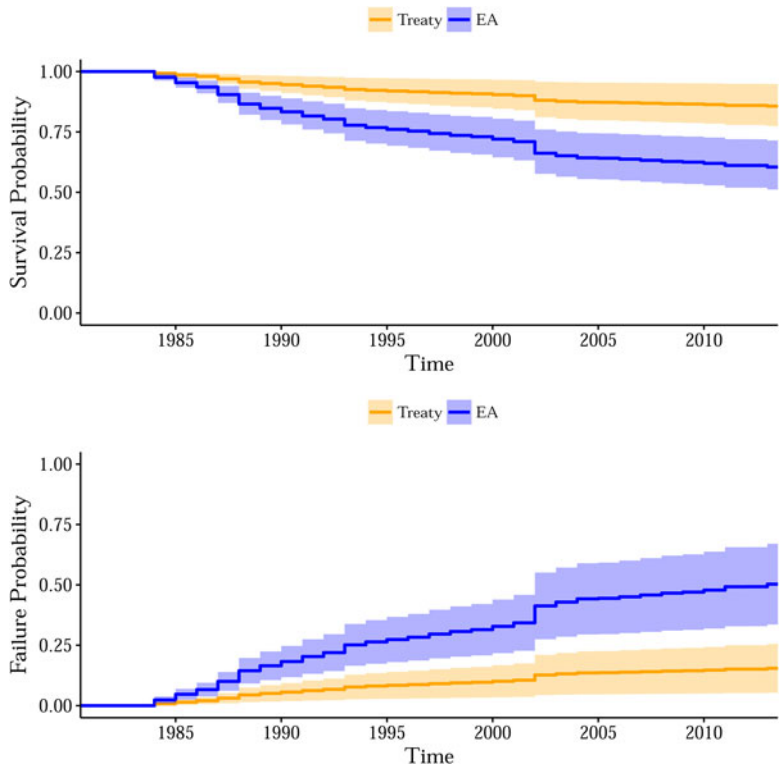

FIGURE 3. Survival and Hazard Curves by Agreement Type

This figure depicts estimated survival curves (top) and estimates hazard curves (bottom) for treaties and agreements over the period of observation. Shaded areas are 95% confidence intervals.

difference in the durability between *ex post* congressional-executive agreements and treaties, that difference is substantively small and statistically insignificant. In every model specification that includes other executive agreements, the difference is large and statistically significant. As such, the evidence is consistent with the view that differences between treaties and congressional-executive agreements are less pronounced for *ex post* congressional-executive agreements.

That said, one cannot conclude that there is definitively no difference between treaties and *ex post* congressional-executive agreements. The failure to reject the null hypothesis is different from proving the null hypothesis. The number of *ex post* congressional-executive agreements in the sample is small. As such, failure to reject the null hypothesis may simply be a result of large standard errors due to data scarcity. This is especially true for Models (2)–(5), which include a large number of covariates and thus lead to data sparsity within many subgroups. The fact that almost all model specifications yield negative coefficients certainly makes it possible that more data may yield a statistically significant difference, albeit a small one.

Consider now the case of sole executive agreements. As pointed out above, TIF does not distinguish between sole and congressional-executive agreements, though the estimated share of the former ranges from 5 percent to 6 percent of all agreements. To accommodate the fact

https://doi.org/10.1017/ajil.2018.103 Published online by Cambridge University Press

TABLE 6.
COX MODEL BY AGREEMENT TYPE

	Executive Agreements	
	Ex Post	Others
Model (1)	−0.716	−1.184***
	(0.693)	(0.355)
Model (2)	0.602	−1.192**
	(0.779)	(0.363)
Model (3)	−0.176	−1.019**
	(1.320)	(0.371)
Model (4)	−0.302	−1.051**
	(1.358)	(0.371)
Model (5)	−0.229	−1.052**
	(1.338)	(0.371)

Note: *p < 0.05; **p < 0.01; ***p < 0.001

*The table depicts the coefficient on the treaty indicator for different model specifications. "*Ex Post*" compares treaties to* ex post *congressional-executive agreements. "Others" compares treaties to the remaining executive agreements.*

that some international instruments may be sole executive agreements that should be excluded from the analysis, this study employs a sensitivity analysis.[102]

The analysis sorts agreements by their durability and assumes that the x quantile are sole executive agreements, where $x \in [0, 0.1]$. For instance, $x = 0.03$ assumes that the 3 percent least durable agreements are sole executive agreements. It then omits these agreements from the analysis, runs the preferred Model (5) and collects the estimated coefficient on the treaty indicator and its standard error. Note that the assumption that the least durable agreements are sole executive agreements is extremely restrictive. In reality, it is much more likely that some sole executive agreements outlast congressional-executive agreements. It can thus be expected that this approach biases the durability of congressional-executive agreements upward, making it harder to detect a difference between the durability of treaties and executive agreements. If it can be shown that even under these restrictive assumptions treaties survive executive agreements, this provides particularly strong evidence for the greater durability of treaties.

Figure 4 reports the estimated coefficients and 95 percent confidence intervals for all x over the range [0, 0.1]. Even under the strict assumption that the 10 percent shortest-lasting executive agreements are sole executive agreements, there still is a substantial difference between treaties and congressional-executive agreements that is statistically different from 0.

V. DISCUSSION AND IMPLICATIONS

This study set out to investigate whether promises made in the form of a treaty are significantly more durable than those made as congressional-executive agreements. The analysis suggests that this in fact is the case. When holding all observed characteristics constant, an

[102] For similar approaches in the sensitivity analysis of causal estimates, see Liu Weiwei, S. Janet Kuramoto & Elizabeth A. Stuart, *An Introduction to Sensitivity Analysis for Unobserved Confounding in Non-Experimental Prevention Research*, 14 PREVENTION SCI. 570 (2013).

https://doi.org/10.1017/ajil.2018.103 Published online by Cambridge University Press

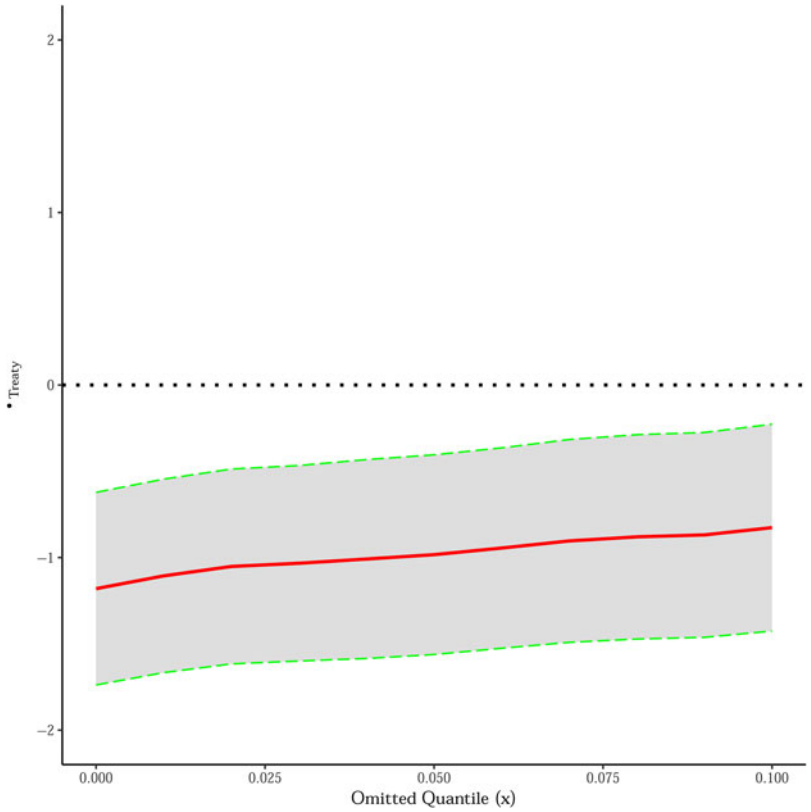

FIGURE 4. Omitting Sole Executive Agreements

This figure depicts the coefficient on the treaty indicator of Model (5) under the assumption that the x quantile of agreements are sole executive agreements and should thus be omitted from the analysis. The shaded area is the 95% confidence interval.

agreement adopted as a treaty lasts statistically and substantively longer than a similar agreement concluded as an executive agreement. This finding holds true for all model specifications and even under the assumption that the 10 percent shortest lasting executive agreements are sole executive agreements that could not easily replace the treaty.

It is important, however, to acknowledge the limitations of this study, which highlight that the above findings should be viewed not as an end point but rather as an important step in understanding the relevance of the choice among international instruments.

The first limitation relates to the causal interpretation of the findings. The choice between treaties and congressional-executive agreement is not random, and there is no guarantee that the estimates derived from the analysis can be interpreted as causal estimates, i.e. that agreements concluded as treaties last long *because* they have been adopted as treaties and not congressional-executive agreements.[103]

[103] Formally, it is required that the covariate of interest, here the treaty indicator, is uncorrelated with the potential outcome, here an agreement's durability, after the inclusion of all control variables. This is also referred to as the assumption of "selection on observables." If the choice between treaties and congressional-executive agreements was random, the covariate of interest would, by definition, be uncorrelated to the potential outcome. However,

https://doi.org/10.1017/ajil.2018.103 Published online by Cambridge University Press

Through the inclusion of several controls and fixed effects, the analysis attempts to address most of the plausible sources of omitted variable bias that are motivated by the different theories, such as selective senatorial attention or subject-specific conventions. The fact that the coefficients on *Treaty* do not vary much even after the inclusion of these controls is a reassuring indication that the results may not simply be the consequence of unobserved selection effects.

However, it cannot be ruled out that at least part of the difference in durability is driven by nuanced considerations that the model is unable to capture. For instance, it is possible that, within a given subject matter, senatorial attention is selectively directed at certain types of agreements, such as those of particular importance, and that this selection is also correlated with durability. The included subject categories may be too crude to capture this dynamic and it is possible that a more granular measure of the subject matter would yield a richer and more complete understanding of the association between the substance of an agreement and the choice of commitment device that may explain at least part of the observed difference.[104]

Future research could address this limitation by sorting agreements on a more granular level than is possible when relying on data from TIF. One particularly promising approach could result from a detailed analysis of the content of individual agreements. Analyzing the text of the agreements would allow researchers to distinguish, for instance, between bilateral tax agreements intended to avoid double taxation, which are often concluded as treaties, and other types of tax agreements. The relevant databases to conduct these types of studies exist,[105] but it has previously been infeasible to read and coherently categorize several thousand international agreements. Recent advances in computer-assisted text analysis could prove fruitful in overcoming this limitation. Indeed, scholars have already begun to employ text analysis in order to evaluate a limited set of international agreements, such as preferential trade agreements[106] or bilateral investment treaties,[107] and the same methodology may be used to examine international agreements on a larger scale. Particularly promising methods include topic modeling and clustering, which allow the researcher to define an arbitrarily large number of categories in which to group a document based on its textual content. This method would enable the automated, granular categorization of international agreements. A more nuanced categorization acquired through such an analysis could further inform our understanding both of the role of the treaty and the possible consequences that abandoning the treaty would have on the landscape of U.S. international agreements.

A second limitation of this study is that, while it suggests that treaties retain value as a policy tool, it does not directly speak to the relative importance of the different hypotheses

absent randomization, the selection on observables assumption remains unverifiable and subject to theoretical debate.

[104] Recall, however, that including subject matter fixed effects does not cause large changes in the coefficients. In order for more nuanced selection effects to fully explain the results, one would have to assume that the within-category selection effects are stronger than the between-category selection effects, for example, that the average difference between two tax agreements explains more variation than the average difference between a tax agreement and an arms limitation agreement.

[105] For instance, HeinOnline's U.S. Treaties and Agreements Library offers access to the full text of a large number of international agreements.

[106] Wolfgang Alschner, Julia Seiermann & Dmitriy Skougarevskiy, *Text of Trade Agreements (ToTA)—A Structured Corpus for the Text as Data Analysis of Preferential Trade Agreements*, 15 J. Empirical Legal Stud. 648 (2018) (describing the creation and subsequent analysis of a database of 448 preferential trade agreements).

[107] Michael Waibel, *Fair and Equitable Treatment as Boilerplate* (not yet published, 2018) (examining the text of "fair and equitable treatment" clauses in investment treaties).

https://doi.org/10.1017/ajil.2018.103 Published online by Cambridge University Press

for the greater durability of treaties. Several mechanisms have been proposed that could potentially explain this durability, ranging from signaling theory, to the stability of senatorial preferences, to the possibility that the advice and consent process may reveal more credible information to negotiating partners. To be sure, none of these explanations are mutually exclusive; indeed, it may be naïve to assume that a single theory can explain the choice between commitment devices for every agreement. However, since all mechanisms in the present analysis lead to observationally equivalent outcomes, the findings provide little guidance to those interested in assessing and comparing the relative importance of each of the proposed explanations.

Notwithstanding these limitations, this study helps to inform the debate surrounding the continuing relevance of the treaty. The findings are consistent with the view that the treaty confers certain benefits on the parties that the congressional-executive agreement does not, in turn leading to agreements that are qualitatively different. Despite the decline in its use, the treaty appears to retain an important role as a policy tool. In particular, the optimal choice among international agreements may require a presidential administration to carefully consider the strength of the commitment, the private information revealed to the public, the domestic audience costs, and the ease with which an agreement can be terminated. Treaties and congressional-executive agreements reflect different tradeoffs among these characteristics, and abandoning the treaty may thus negatively affect the executive's ability to tailor an agreement to a particular context. Consequently, policy recommendations calling for the abandonment of the treaty seem premature and may result in unintended consequences.

To illustrate, consider again the negotiations surrounding SALT II and SORT, where Russia insisted on the use of a treaty over a congressional-executive agreement. Without the availability of the treaty instrument, it is conceivable that the parties would have reached agreements with substantively different terms. Given that Russia would not have spent any of its bargaining power on the agreement type and could instead have devoted it fully to the substance of the agreements, it appears at least plausible to assume that these alternative terms may have been more favorable to Russia. Under the assumption that the counterparties' desire for an instrument with the characteristics of the treaty is strong enough, it may even be the case that, absent the treaty, certain agreements would not have come to fruition at all.

As noted above, some of the consequences of abandoning the treaty may be mitigated by choosing *ex post* congressional-executive agreements.[108] This Article's findings cannot rule out this possibility, given that data sparsity precludes comparing the durability of treaties and *ex post* congressional-executive agreements with much confidence. With that said, this study's review of all known statutory authorizations between 1980 and 2000 at least raises serious concerns to that end.

Indeed, the dynamics surrounding the *ex post* congressional-executive agreement are often described as if the executive submitted to Congress a specific agreement and Congress then considered that agreement in isolation. This is certainly an accurate description for legislative acts implementing important trade agreements. However, outside of the area of trade, the approval process often looks quite different. There, the review of known approval legislation suggests that (1) the language is often vague and it is unclear whether it is approving any specific

[108] *See* Hathaway, *supra* note 1, at 1307; *see also* Oona A. Hathaway, *Presidential Power Over International Law: Restoring the Balance*, 119 YALE L.J. 140, 260 (2009).

https://doi.org/10.1017/ajil.2018.103 Published online by Cambridge University Press

agreement in particular; and (2) approval is typically only one small aspect in an act addressing a much broader set of substantive issues. To illustrate, consider the following language in the 1996 Balanced Budget Downpayment Act, which is thought to approve the Global Learning and Observations to Benefit the Environment (GLOBE) agreement:[109]

> [T]he rate for operations only for program administration and the continuation of grants awarded in fiscal year 1995 and prior years of the Advanced Technology Program of the National Institute of Standards and Technology, and the rate for operations for the Ounce of Prevention Council, Drug Courts, Global Learning and Observations to Benefit the Environment, and for the Cops on the Beat Program may be increased up to a level of 75 per centum of the final fiscal year 1995 appropriated amount.[110]

While the language implies that Congress approves of the GLOBE agreement, it is difficult to read an explicit authorization into the statute. In addition, the reproduced part of the act is the only time GLOBE is mentioned and includes a total of ninety-seven words. The entire act, however, is over 10,000 words long and got approved in a single roll-call vote both in the House and the Senate.[111] So even if one were to read this as an *ex post* authorization of GLOBE, the authorizing text would result in less than 1 percent of the total text of the statute. Outside of trade, provisions such as these, where *ex post* congressional-executive agreements are ostensibly approved as one small part of a larger legislative package, are the rule, rather than the exception.[112] This leads to an approval process that is strikingly different from the advice and consent process for treaties. The latter focuses the entire and undivided senatorial attention on the approval of the agreement itself and does not directly tie its success to the future of other policy implementations. It is thus unclear whether *ex post* congressional-executive agreements could consistently provide the same benefits that the treaty grants. As pointed out above, empirical analyses may provide an informative answer to this question only after the *ex post* congressional-executive agreement is observed more frequently and in a wider variety of subject areas.

In addition to providing an affirmation of the importance of the treaty as a policy tool, this Article's findings also have implications for current debates surrounding the presidential power to withdraw from international agreements. Under the Trump administration, the United States has announced its withdrawal from a number of highly publicized

[109] *See* Hathaway, *supra* note 1, at note 49.

[110] Balanced Budget Downpayment Act, Pub. L. No. 104-99, § 201(a), 110 Stat. 26, 34 (1996).

[111] There were a few attempts in the Senate to waive the act which were rejected.

[112] Of the five legislative acts not related to trade, only one is an implementation act. The others address a larger set of policy goals. In addition to the 1996 Balanced Budget Downpayment Act, see Atomic Energy Act – Exemption, Pub. L. No. 109-401, 120 Stat. 2726 (2006) (including detailed provision which waive certain requirements of the Atomic Energy Act of 1954, and map out the legal framework for a proposed, future agreement between the United States and India); South African Democratic Transition Support Act of 1993, Pub. L. No. 103-149, 107 Stat. 1503 (1993) (a general act setting out United States policy with respect to the South African transition process after Apartheid); Support for East European Democracy (SEED) Act of 1989, Pub. L. No. 101-179, 103 Stat. 1298 (1989) (a general act including many provision designed "[t]o promote political democracy and economic pluralism in Poland and Hungary . . ."). The exception is the Agreement for Cooperation Concerning Peaceful Uses of Nuclear Energy, Pub. L. No. 99-183, 99 Stat. 1174 (1985) (codified at 42 U.S.C. § 2156 (2000)) (authorizing a prior agreement between the United States and China on the peaceful use of nuclear energy).

https://doi.org/10.1017/ajil.2018.103 Published online by Cambridge University Press

agreements—such as the Paris Agreement on Climate Change, the Joint Comprehensive Plan of Action with Iran, the 1955 Treaty of Amity with Iran, and the Optional Protocol to the Vienna Convention on Diplomatic Relations—and threatened to withdraw the United States from NAFTA and NATO.

There is an active scholarly debate surrounding the scope of presidential power to withdraw the United States from its international agreements, with a particular focus on whether the constitutional hurdles to terminate congressional-executive agreements are higher than those for the treaty.[113] The results of this study suggest that, in addition to the doctrinal issues, a political economy analysis may provide valuable insights as well.

Indeed, the finding that treaties outlast congressional-executive agreements could be interpreted as indicating that treaties are harder to terminate, if not as a matter of law, then at least as a matter of political reality. We currently lack a comprehensive theory of the political costs of treaty termination. To be sure, the writings of some commentators imply that part of the political cost differential between breaching a treaty and breaching a congressional-executive agreement may be found in reputational sanctions. That is, the use of a treaty would "represent the complete pledge of a nation's reputational capital"[114] and presidents that use it "somehow put it all on the line in the diplomatic world."[115] However, in order for this mechanism to explain why treaties are not terminated at the same rate as congressional-executive agreements, one would have to assume that the United States has a single reputation that is independent of the administration currently in power. Otherwise, it would be difficult to understand why future administrations should feel bound to promises their predecessors made.[116]

This study's empirical findings are ill-suited to evaluating the theoretical merits of these claims. Future research aimed at understanding not only the legal differences but also the political implications of withdrawing from a treaty *vis-à-vis* the congressional-executive agreement could provide further insights and shed light onto the question of whether the fate of international agreements such as the Paris Agreement or NAFTA would have been less controversial had they been concluded as treaties instead.

VI. Conclusion

Relying on survival time analysis, this Article reveals that treaties are more durable commitments than executive agreements. In particular, there was a 15 percent probability that a typical agreement concluded as a treaty in 1982 broke down by 2012, compared to a 50 percent probability that it broke down when concluded as an executive agreement. The findings are consistent with the view that treaties remain an important policy tool for the United States because they represent a promise that is qualitatively different from a promise made as an *ex ante* congressional-executive agreement.

[113] *See* Bradley, *supra* note 36; Galbraith, *supra* note 26; Curtis A. Bradley & Laurence R. Helfer, *Treaty Exit in the United States: Insights from the United Kingdom or South Africa?*, 111 AJIL Unbound 428 (2017).

[114] Andrew T. Guzman, *A Compliance-Based Theory of International Law*, 90 Calif. L. Rev. 1823, 1880 (2002).

[115] Beth A. Simmons, Mobilizing for Human Rights: International Law in Domestic Politics 120 (2009). *See* also Andrew T. Guzman & Timothy L. Meyer, *International Soft Law*, 2 J. Legal Analysis 171 (2010).

[116] Some scholars are thus critical of the potential for reputational mechanisms to explain compliance. *See, e.g.*, Rachel Brewster, *Unpacking the State's Reputation*, 50 Harv. Int'l L.J. 231, 249 (2009).

https://doi.org/10.1017/ajil.2018.103 Published online by Cambridge University Press

At the same time, the Article raises new questions regarding the mechanism responsible for the greater durability of treaties. The empirical findings suggest that renewed scholarly focus on this issue, and on the political costs of agreement termination, provide fruitful ground in which to grow our understanding of the practical implications for U.S. policy of choosing between treaties and congressional-executive agreements.

Appendix

Details on Model Selection

As explained in the text, because international agreements can go out of force at any time, survival times are continuous in nature. However, since survival times are measured only once per year when TIF is published, the data can best be described as continuous data that is grouped by year. For truly continuous data in which an event can happen at any point in time, the Cox proportional hazard model[117] has established itself as the preferred choice by researchers,[118] as it is a semi-parametric model that only relies on few assumptions. The popularity of this model stems from the fact that it can be estimated without making any parametric assumptions about the baseline hazard rate. For instance, the researcher does not need to assume that survivability decreases at a constant rate over time, exponentially or in any other predefined way. However, the Cox model assumes that there are no ties in the data, meaning that no two observations have the exact same survival time. This is due to the fact that ties cannot occur if survival times are measured on a truly continuous scale. Researchers have developed several techniques make to the Cox model tractable even in the presence of ties. The most precise approach is the "exact method" developed by Kalbfleisch and Prentice.[119] Intuitively, if two subjects i and k survive exactly n periods, the exact method considers the alternative that i survived longer than k and the alternative that k survived longer than i and opts for the one that is more likely.[120] However, in datasets with many subjects, periods, and ties, the exact method is not feasible as it is computationally very intensive. The "Efron method"[121] provides an approximation to the exact method that does not suffer from comparable resource constraints but is a little less precise.

An alternative to the Cox model is a parametric survival model, meaning a model in which the baseline hazard rate is assumed to be of a specific form. Among the parametric models, the

[117] David R. Cox, *Regression Models and Life Tables*, 34 J. Royal Stat. Soc'y 187 (1972).

[118] In general, *see* Danyu Y. Lin, L. J. Wei & Z. Ying, *Checking the Cox Model with Cumulative Sums of Martingale-Based Residuals*, 80 Biometrika 557, 557 (1993) ("The proportional hazards model with the partial likelihood principle . . . has become exceedingly popular for the analysis of failure time observations.") (citations omitted); *see also* Tian, Zucker & Wei, *supra* note 91, at 172 ("The most popular semiparametric regression model for analyzing survival data is the proportional hazards (PH) model.") (citation omitted). For examples in international law, *see* Elkins, Guzman & Simmons, *supra* note 91, at 828 (estimating adoption times for bilateral investment treaties using a Cox model); Simmons, *supra* note 91, at 823 (2000) (relying on the Cox model to estimate time until states accept commitments under IMF Articles of Agreement Article VIII).

[119] John D. Kalbfleisch & Ross L. Prentice, *Marginal Likelihoods Based on Cox's Regression and Life Model*, 60 Biometrika 267 (1973).

[120] Through maximization of the associated likelihood function.

[121] Bradley Efron, *The Efficiency of Cox's Likelihood Function for Censored Data*, 72 J. Am. Stat. Ass'n 557 (1977).

https://doi.org/10.1017/ajil.2018.103 Published online by Cambridge University Press

complementary log-log discrete model[122] is the uniquely appropriate model for grouped continuous data.[123]

Whether to prefer the Cox model in combination with an Efron approximation over the complementary log-log discrete model cannot be answered in a general way. Simulations show that even with heavily tied datasets, the Efron approximation often achieves very accurate results.[124] As a rule of thumb, Chalita et al. propose to compute the quantity

$$pt = \frac{nf - r}{n}$$

where nf is the number of events (here, agreements that went out of force), r is the number of unique survival times, and n is the number of agreements.[125] For $0 \leq pt < 0.2$, Chalita et al. suggest a continuous model with likelihood approximation; for $0.2 \leq pt \leq 0.25$, both discrete and continuous models can be used; for $pt > 0.25$, a discrete model is preferred. Here, $pt = 0.19$, which is why a Cox proportional hazard model with Efron approximation is used in the primary model specifications. The complementary log-log model serves as a robustness check.

Both the Cox and the complementary log-log model rely on the assumption that the hazard is proportional to the baseline hazard ratio. This assumption can be tested using the Grambsch and Therneau method,[126] which plots the Schoenfeld residuals against the rank

[122] The complementary log-log discrete model (or c-log-log model) is of the form

$$h_i(t|x_i) = 1 - (1 - h_0(t_i))^{e^{x_i'\beta}}$$

or, if linearized,

$$\log(-\log(1 - h(t))) = \alpha_j + x_i'\beta$$

where j denotes grouped time intervals. Note that

$$a_j = \log(-\log(1 - h_0(t_j)))$$

is an interval-specific complementary log-log transformation of the baseline hazard rate, $h_0(t_j)$. This means that the baseline hazard rate is allowed to vary with each interval, thus imposing only mild parametric assumptions.

[123] JOHN D. KALBFLEISCH & ROSS L. PRENTICE, THE STATISTICAL ANALYSIS OF FAILURE TIME DATA 47 (2d ed. 2002). The statement refers to the continuous-time proportional-hazards model, where observations have been grouped by time. McCullagh shows that this model is identical to the complementary log-log discrete model, *see* Peter McCullagh, *Regression Models for Ordinal Data*, 42 J. ROYAL STAT. SOC'Y, SER. B (METHODOLOGICAL) 109 (1980).

[124] *See* Irva Hertz-Picciotto & Beverly Rockhill, *Validity and Efficiency of Approximation Methods for Tied Survival Times in Cox Regression*, 53 BIOMETRICS 1151 (1997); Liciana V.A.S. Chalita, Enrico A. Colosimo & Clarice G. B. Demétrio, *Likelihood Approximations and Discrete Models for Tied Survival Data*, 31 COMMUNICATIONS IN STATISTICS-THEORY AND METHODS 1215 (2002); Jadwiga Borucka, *Methods of Handling Tied Events in the Cox Proportional Hazard Model*, 2 STUDIA OECONOMICA POSNANIENSIA 91 (2014).

[125] Chalita, Colosimo & Demétrio, *supra* note 124, at 1220.

[126] Patricia M. Grambsch & Terry M. Therneau, *Proportional Hazards Tests and Diagnostics Based on Weighted Residuals*, 81 BIOMETRIKA 515 (1994).

https://doi.org/10.1017/ajil.2018.103 Published online by Cambridge University Press

of the time intervals. If the proportionality assumption holds, then there should be no systematic pattern. In a formal test of non-proportionality, thirty-one of 264 (or 11 percent) covariates yield significant p-values implying a violation of non-proportionality. Reassuringly, the covariates of interest are not among them. However, even for the remaining covariates, the disproportionality is of little concern for two reasons.

First, note that the Grambsch and Therneau test was developed in the medical context where sample sizes are typically smaller than one hundred, making the test insensitive to minor disproportionalities. For sample sizes as large as in this study, small confidence intervals lead to significant p-values even if the data reveals negligible disproportionalities. In addition to the formal tests, visual examination of the Schoenfeld residuals is thus recommended.[127] Such a visual examination yields no significant violations of the proportional hazards assumption for any of the subject matter covariates, and a violation only for a handful of countries, typically those with whom the U.S. has only few agreements, such as Burma, Ecuador, or New Caledonia. The corresponding graphs are included in the online appendix.

Second, note that the concern for a violation of the proportional hazard assumption stems from the medical sciences, where it is of great importance whether a drug has an inverse or possibly a reverse effect on a subgroup of patients. However, in the social sciences, researchers are typically interested in average covariate effects across the entire sample. As Allison highlights, even in cases where the proportionality assumption is violated, estimates can still be interpreted as average covariate effects.[128] Violations of the proportionality assumption thus do not present a threat to the interpretability of the coefficients for most social scientific studies such as the present one.[129]

Online Appendix

Supplementary material for this article can be found at https://doi.org/10.1017/ajil.2018.103.

[127] Eric Vittinghoff, David V. Glidden, Stephen C. Shiboski & Charles E. McCulloch, Regression Methods in Biostatistics: Linear, Logistic, Survival, and Repeated Measures Models 237 (2d ed. 2012) ("The Schoenfeld test is widely used and gives two easily interpretable numbers that quantify the violation of the proportional hazards assumption. However, . . . in large samples they may find statistically significant evidence of model violations which do not meaningfully change the conclusions.").

[128] Paul D. Allison, Survival Analysis Using SAS: A Practical Guide 173 (2d ed. 2010) (pointing out that interactions with time are commonly suppressed and that the estimates are nonetheless meaningful averages); *see also* Paul Allison, Event History and Survival Analysis 43 (2d ed. 2014) ("Even when the proportional hazards assumption is violated, it is often a satisfactory approximation. Those who are concerned about misspecification would often do better to focus on the possibilities of omitted explanatory variables, measurement error in the explanatory variables, and informative censoring.").

[129] Indeed, such a scenario is comparable to the process of fitting a linear regression model to non-linear data. The reason why the OLS regression is so popular in many social scientific applications is that the obtained coefficients can still reasonability be interpreted as average covariate effects, even though the data generating process is non-linear. That is why the linearity assumption is hardly ever validated.