

SYMPOSIUM ON JULIAN NYARKO, “GIVING THE TREATY A PURPOSE: COMPARING THE DURABILITY OF TREATIES AND EXECUTIVE AGREEMENTS”

DURATION ≠ SERIOUSNESS OF COMMITMENT: AN EMPIRICAL AND THEORETICAL CRITIQUE OF NYARKO’S TREATIES VS. EXECUTIVE AGREEMENTS

Barbara Koremenos*

In “Giving the Treaty a Purpose,”¹ Julian Nyarko distinguishes between treaties and executive agreements and argues that treaties signal a higher² level of commitment to our partners in cooperation than do executive agreements because treaties are more durable. Nyarko uses survival-time analysis to demonstrate that treaties last longer than executive agreements—that is, treaties are less likely to drop out of the *Treaties in Force* (TIF) series in any given year. The longer life of treaties is Nyarko’s proxy for their greater durability. Nyarko argues that his result holds “even after controlling for a number of covariates that could influence the durability of the agreement,” like particular presidents, subject areas, and partner countries as well as the degree of divided government. Nonetheless, Nyarko’s list omits the most important variable affecting durability as he defines it: intended duration. Sometimes the intended duration of a piece of formal international law is finite. Indeed, as I will explain in this response, under certain (and common) conditions, *this choice of a finite duration is what makes the commitment credible* (or, in Nyarko’s language, reliable).

Nyarko uses both the jargon of failure/death and the statistical methods most often employed in medicine regarding time to death for terminally ill patients. But to describe an agreement that appears in TIF one year and not the next as a failure/death ignores the possibility that the parties never intended the agreement to last forever. In fact, according to the Continent of International Law (COIL) dataset (described below), more often than not international agreements are *designed to be finite*.³ Moreover, many of these finite-duration agreements are followed up by another finite-duration agreement or even a series of finite agreements.

Finite Duration by Design: A Few Examples

Consider the many bilateral U.S. agricultural commodity agreements designed to give food aid to a developing country and allow farmers to sell surplus commodities, especially under conditions of depressed market prices.⁴ Such agreements constitute a nontrivial set of the executive agreements (EAs) used in Nyarko’s

* Professor, Department of Political Science, University of Michigan.

¹ Julian Nyarko, *Giving the Treaty a Purpose: Comparing the Durability of Treaties and Executive Agreements*, 113 AJIL 54 (2019).

² Nyarko also uses the words “more serious.”

³ BARBARA KOREMENOS, *THE CONTINENT OF INTERNATIONAL LAW: EXPLAINING AGREEMENT DESIGN* (2016).

⁴ See, e.g., Agricultural Commodities Agreement Between the United States of America and the Government of the Republic of China Under Title I of the Agricultural Trade Development and Assistance Act (with Exchange of Notes), as Amended, Aug. 30, 1960, 388 UNTS 5579.

analysis.⁵ These are ex ante congressional-executive agreements,⁶ with many of them implementing Title I of the Agricultural Trade Development and Assistance Act.⁷

Originally, the Act and the bilateral agreements it provided for were intended to establish a market for surplus U.S. produce⁸ and to last only three years.⁹ Over time, the Act's policy objective encompassed not only the stabilization of domestic market prices and the development of export markets,¹⁰ but also humanitarian aid.¹¹ The agreements implementing the Act were thus time-limited; the content of these agreements—for example, the kinds and quantities of produce exported by the United States—also changed.¹²

The United States and Taiwan concluded four such EAs between 1957 and 1967. The first lists cotton, dairy products, tobacco, and inedible tallow as the approved commodities,¹³ but tobacco and tallow are removed in some of the subsequent agreements.¹⁴ The total export market value in each agreement also fluctuates, increasing in the second agreement, decreasing in the third, and increasing again in the fourth.¹⁵ Cotton is a significant export in all four agreements. Another noticeable trend is the increasing precision in agreement conditions over time, with the fourth agreement adding specific supply periods for the commodities, limits on maximum quantities of exports, and requirements for shipping payments. This set of agreements demonstrates that the cooperation between the parties is quite durable, despite the fact that any particular agreement governing that cooperation may be finite. Both parties benefit when they can adjust the specifics every few years based on supply and demand conditions; a static set of terms is not economically optimal.¹⁶ In assessing reliability, our focus should be the *durability of the cooperative endeavor*, not the durability of any single agreement in this cooperative endeavor.

The finite-duration 1962 International Coffee Agreement provides another case in point. This agreement stipulated a quota system for coffee exports to help raise and stabilize the price of coffee and called on coffee importers to police and enforce the quota system. Without the finite-duration provision, the battle over the exporters' quota shares, which also determined voting weights, would have been intractable. South American states, such as Brazil and Colombia, wanted to lock in their dominance at the time, whereas many African states, which were witnessing rapid growth in their production capabilities as well as increased demand for their coffee, wanted a

⁵ See Nyarko, *supra* note 1, at Table 2.

⁶ See, e.g., Curtis A. Bradley, *Exiting Congressional-Executive Agreements*, 67 DUKE L.J. 1616, 1644 (2018).

⁷ *Agricultural Trade Development and Assistance Act of 1954*, ch. 469, Pub. L. No. 480, 68 Stat. 454. Similarly, many of the security agreements in the COIL random sample (labeled as “defense” under TIF) implement the *Mutual Security Act of 1951*, Pub. L. No. 82-105, 65 Stat. 373.

⁸ H.R. REP. NO. 1776, *as reprinted in* 1954 U.S.C.C.A.N. 2509.

⁹ JOHN R. TARRANT, *FOOD POLICIES* 237 (1980).

¹⁰ See SENATE COMM. ON AGRICULTURE, NUTRITION, AND FORESTRY, 98TH CONG., 2D SESS., *TRADE POLICY PERSPECTIVES: SETTING THE STAGE FOR 1985 AGRICULTURAL LEGISLATION* 280 (Comm. Print 1984).

¹¹ See *id.* at 225.

¹² James R. Walczak, *New Directions in United States Food Aid: Human Rights and Economic Development*, 8 DEN. J. INT'L L. & POL'Y 543, 548 (1979).

¹³ *Agricultural Commodities Agreement Under Title I of the Agricultural Trade Development and Assistance Act (with Exchange of Notes)*, Aug. 14, 1956, 281 UNTS 4083.

¹⁴ *Agricultural Commodities Agreement Under Title I of the Agricultural Trade Development and Assistance Act, as amended (with Exchange of Notes)*, Aug. 30, 1960, UNTS 5579; *Agricultural Commodities Agreement Under Title I of the Agricultural Trade Development and Assistance Act, as amended (with Exchange of Notes)*, June 3, 1964, UNTS 7610.

¹⁵ *Agreement for Sales of Agricultural Commodities (with Annex)*, Dec. 12, 1967, UNTS 10061.

¹⁶ The United States can also leverage the promise of a continued relationship or the threat of a terminated one to achieve its foreign policy objectives in the country receiving the aid.

flexible approach. In the end, the *limited duration and renegotiation provision* of the 1962 agreement provided the solution: the 1962 International Coffee Agreement could be sustained as an equilibrium institution because it was in the interest of all the states involved to adhere to it for its limited duration. Once it came to an end, cooperation in coffee could continue as an equilibrium outcome because states could negotiate a new agreement to reflect the new economic and political realities. Indeed, the 1962 distribution of quotas was altered in 1968, 1976, and 1983.¹⁷ Thus, like the agricultural commodity agreements discussed above, the set of finite-duration coffee agreements facilitated a durable cooperative endeavor.

It is also quite likely that U.S. participation in the agreement would not have been forthcoming if it had been given an indefinite duration. U.S. motives were geopolitical.¹⁸ The United States feared the economies of its Latin American neighbors were in danger of collapse because of dramatic instability in the price of coffee. Hence, the United States was willing to depart from its free-market principles in the short run to keep these countries on its side of the Cold War divide.

Importantly, once coffee prices increased and stabilized, the Latin American economies diversified, and the United States won the Cold War, the coffee cartel was no longer necessary—that is, the problems that dictated the cartel were solved.¹⁹ An apt medical analogy is a treatment for the multitude of medical problems that are of a short-to-medium term and solvable nature. A broken limb and even multiple types of cancers necessitate visits to the hospital and treatment. When the treatment and hospital visits stop because the limb has healed or the cancer is definitively removed, we call this success, not failure. Put differently, a reliable treatment may be short-term if that is what is needed to solve the problem. The same is true regarding a reliable international agreement.

Choosing the Optimal Duration

How can one determine whether a finite agreement would make cooperation more reliable? The (un)predictability of the underlying environment is key. States often have large amounts of information at their disposal when initiating cooperative activity and use it wisely to set the terms of cooperation. Inevitably, though, events occur that could not have been foreseen and cannot be controlled. These events often alter the gains and losses from cooperation. The likelihood of such events (we can call them exogenous shocks) varies across subissue areas, with some, like agricultural commodities, being far more vulnerable than others, say the human rights issue area. This uncertainty regarding the consequences of cooperation is what the COIL framework calls *Uncertainty About the State of the World* (U(SofW)).²⁰

The uncertainty can be scientific and technical or it can be about politics or economics. Even in the absence of shocks, there can still be uncertainty about how an agreement will work in practice, especially in more complex issue areas. This is akin to an “experience good” in economics: the true nature of the good can only be known with experience (that is, time).

One of the hypotheses highlighted in the COIL research program is as follows: other things equal, when the underlying U(SofW) is high for a particular cooperative endeavor, finite-duration agreements are optimal. Case study evidence, including from the Nuclear Nonproliferation Treaty negotiations, illustrates the causal mechanism the theory articulates. Large-n testing strongly supports the hypothesis. Let me elaborate.

¹⁷ These changes were not trivial. For example, between the 1962 and 1983 agreements, Brazil's share of the total export quota fell from 39.2% to 30.8%. See Barbara Koremenos, *Can Cooperation Survive Changes in Bargaining Power? The Case of Coffee*, 31 J. LEGAL STUD. 259, 259–83 (2002).

¹⁸ See *International Coffee Agreement: Hearing Before the H. Comm. on Ways and Means*, 89th Cong. 5, 9 (1965).

¹⁹ In the case of coffee, the problems were solved over a twenty-five-year period by a series of agreements.

²⁰ See KOREMENOS, *supra* note 3.

The empirical portion of COIL includes a data set featuring 234 randomly selected agreements drawn from the United Nations Treaty Series (UNTS) across the issue areas of economics, environment, human rights, and security. Two separate sets of coders for the cooperation problems (the independent variables) and the hundreds of design dimensions (the dependent variables) were employed to preserve the integrity of the project. An important part of the research program features the careful definition and operationalization of the underlying cooperation problems, like U(SofW), so that they can be identified across the sample as either present in a high degree or low degree. For example, economic agreements subject to supply-and-demand shocks like those governing coffee or sugar quotas meet the threshold for high U(SofW).²¹

Nyarko's dataset consists of the population of treaties reported in TIF that were signed and ratified by the United States between the years 1982–2012. The COIL dataset is a random sample covering the entire world and a much longer time period: 1925–2004, with over two-thirds of the sample consisting of agreements concluded between 1970 and 1999. The COIL sample draws from the following UNTS subject areas: agricultural commodities, disarmament, environment, finance, human rights, investment, monetary matters, and security (labeled as defense by TIF). The majority of subject areas (62 percent) represented in Nyarko's analysis are represented in COIL, including agriculture, arms limitation, education, postal matters, defense, labor, finance, trade and commerce, environment, fisheries, human and fundamental rights, maritime matters, and taxation. Furthermore, given that Nyarko's theory is about signaling commitment to our partners in cooperation, the fact that the COIL dataset excludes bilateral agreements between one state party and an international organization, whereas TIF does not, is not a weakness.²²

While the United States is the most represented state in the COIL sample, there are only seventy-six agreements in which the United States participated.²³ Still, in work using the COIL data set, there has been enough power in the sample to make inferences—that is, statistically significant results are obtained and marginal effects are impressive.²⁴ Tables 1 and 2 describe the portion of the COIL sample characterized by U.S. participation; these are the agreements used in the analyses depicted in Tables 3 and 4.²⁵

Controlling for the underlying environment is critical in any analysis of how duration affects the reliability of commitment. In the COIL sample, 60 percent of the agreements are characterized by an underlying U(SofW). Table 3 presents some descriptive statistics on the seventy-six agreements featuring the United States and shows EAs are far more likely to be characterized by an underlying U(SofW) than are treaties. And, not surprisingly, these EAs are far more likely to be designed to be finite, as illustrated in Table 4.

²¹ For this problem to meet the “high” threshold, coders had to argue that the environment was one in which changes or shocks could cause the distribution of gains from the agreement to vary substantially over time or in which great uncertainty existed about how the agreement would work in practice. This is elaborated in Barbara Koremenos, *Contracting Around International Uncertainty*, 99 AM. POL. SCI. REV. 549 (2005).

²² An example of an agreement that would be excluded from COIL on this basis is the [Fourth Supplemental Agreement Regarding the Headquarters of the United Nations, with annex](#), June 18, 2009, TIAS 09-618.

²³ Almost all of the major human rights and arms control agreements are in the COIL sample as are double-taxation agreements, bilateral investment treaties, major multilateral environmental agreements, and the Geneva Convention Relative to the Treatment of Prisoners of War.

²⁴ COIL does not attempt to approach the population of UNTS agreements. Still, although more power (in the statistical sense) is always preferred to less, and the larger the sample size, the larger the power, power does not increase linearly with sample size. Marginal increases in sample size have a decreasing marginal effect on the variance of estimators. For COIL, the judgment was made to cover a large number of variables and invest heavily in both intercoder reliability and the separation of coders for independent and dependent variables, with the tradeoff of covering fewer agreements.

²⁵ With respect to the “0” in the “Investment” row in Table 2, the Bilateral Investment Treaties in the COIL sample do not happen to include the United States. For a list of the COIL agreements with U.S. participation as well as robustness checks on the analyses in Tables 3 and 4, see the online appendix available from the author (koremeno@umich.edu).

Table 1: Number of EAs and Treaties in COIL Sample

Kind	n
EA	50
Treaty	26

Table 2: Number of EAs and Treaties in COIL Sample by Subissue Area

	EA	Treaty
Agriculture	14	0
Disarmament	1	5
Environment	7	7
Finance	4	2
Human Rights	1	9
Investment	6	0
Security	17	3

Table 3: Presence of U(SofW) by Type of Agreement
(Percentage of Each Type of Agreement)

	Uncertainty World (%)	No Uncertainty World (%)
Treaty	35	65
EA	80	20
<i>p-value</i>	0.0002	
N = 76		

Table 4: Duration Provisions by Type of Agreement
(Percentage of Each Type of Agreement)

	Finite (%)	Indefinite (%)
Treaty	19	81
EA	78	22
<i>p-value</i>	0.0000	
N = 76		

The extremely small p-values from the chi-squared test of the null of independence of the rows and columns demonstrate that, even with the small sample size, there are statistically significant differences between EAs and treaties on these two dimensions.²⁶

Choosing the correct durational provision given the underlying environment is one important means of signaling seriousness of commitment, but it is not the only one. There are other design provisions that signal

²⁶ Put simply, there is an association between the type of agreement and the duration. This result holds even when bootstrapping is performed and a 1% confidence interval is constructed.

commitment that Nyarko also omits from his discussion: provisions calling for delegated dispute resolution, third-party monitoring, and even punishment. Such mechanisms, when designed to confront the underlying cooperation problems, will keep parties from renegeing on their commitments. For example, monitoring mechanisms allow tit-for-tat reciprocity to operate even in the absence of formal punishment mechanisms.²⁷ These types of design provisions occur in both executive agreements and treaties and in agreements of both finite (e.g., superpower arms control) and indefinite duration.

Finally, we cannot equate permanent hands-tying with a deeper commitment. One reason the United States might be willing to ratify a treaty with an indefinite duration is that the treaty is so vague that it commits parties to very little or requires little in terms of changed behavior. This is the case for many of the indefinitely long human rights agreements that the United States has ratified as treaties. Indeed one could argue that the depth of the commitment and the departure from the status quo under the series of finite-duration coffee agreements were much greater than, say, the depth of the commitment and departure from standard operating procedure under the indefinitely long Convention on the Elimination of All Forms of Racial Discrimination (CERD).²⁸

Conclusion

Nyarko fails to control for or even consider theoretically the design feature of an intentionally finite duration. Theory tells us that this institutional design feature is key to the success of international cooperation through formal international law, and empirical testing confirms the theory—in other words, *the limited duration provision often impedes “failure/death.”* To not take this into account in an analysis of the durability of treaties versus executive agreements is a major research design flaw that calls into question the conclusions Nyarko draws.

Hence, while Nyarko’s assertion that treaties play an important role in how the United States cooperates may be correct, his argument and analyses are extraneous to the issue. The COIL theoretical framework and data suggest an alternative explanation: the choice between EAs and treaties may be driven at least in part by the underlying problem structure. If EAs are easily amended or replaced, they are the optimal form of finite-duration agreements in the shadow of uncertainty.²⁹ Future research should examine whether other cooperation problems align with either treaties or EAs and delineate the logic that explains the correspondence.

²⁷ Paul R. Milgrom et al., *The Role of Institutions in the Revival of Trade: The Law Merchant, Private Judges, and the Champagne Fairs*, 2 *ECON. & POL.* 1 (1990).

²⁸ In this example, both international agreements were concluded as treaties. But my point is that the finite coffee treaty (which dropped out of TIF) arguably represents a deeper commitment than the CERD (which remains in TIF).

²⁹ According to the COIL sample, EAs are far more likely to be bilateral whereas treaties are far more likely to be multilateral. Bilateral agreements may be easier to replace than multilateral agreements given that transactions costs generally increase with the number of parties involved, other things equal.