

国际条约约束力 评价体系研究^{*}

赵雪霏

【内容提要】 签署国际条约是国家间合作的重要表现形式,随着各学科的交叉发展,条约的文本设计细节逐渐成为条约约束力研究的焦点问题。基于新制度主义中的理性选择理论,通过对条约文本的考察,可以构建包含成员资格、议题范围、集中度、控制度和灵活度五个维度的条约约束力评价体系,并以中国与东盟国家签署的海洋合作条约作为样本数,利用文本分析和熵值法的计算方法验证该体系。经计算,样本内多边条约的控制度指标和双边条约的灵活度指标被赋予更高权重,符合双边条约与多边条约存在约束差异的假设。同时,样本内多边条约的约束力评价得分均值呈逐年上升、约束差异呈逐年缩小的趋势,双边条约的约束力评价得分均值自2000年后呈逐年下降、约束差异呈逐年缩小的趋势。构建国际条约约束力评价体系,对不同主题、不同国家之间的条约设计特征研究以及对国家签署的同类型条约的约束变化考察都具有一定的应用价值。

【关键词】 国际条约;条约约束力指标;国际海洋合作;“一带一路”;中国—东盟关系

【作者简介】 赵雪霏,厦门大学南洋研究院博士研究生(厦门 邮编:361005)。

【中图分类号】 D815 **【文献标识码】** A **【文章编号】** 1006-9550(2022)10-0130-32

^{*} 本文系中国—东盟海上合作基金年度支持项目“中国—东盟海洋大数据综合信息服务平台”(项目批准号:220203993022761133)后续成果。感谢《世界经济与政治》匿名审稿人的意见与建议,文中错漏由笔者负责。

国家间签署的双边条约或多边条约以文本形式落实,为国际合作提供了一定的法律约束保障,但国际社会的无政府状态又使得一些条约在执行过程中由于违约和实施困难等原因失效。因此,国际条约的有效性——关于硬法(hard law)和软法(soft law)的区分——在国际法和国际关系领域展开了激烈讨论,但在争论的早期阶段这些观点并未对条约的约束力制定出具体的评价规则。^①直到20世纪90年代,随着各学科的交叉发展,一些学者开始将条约约束力的研究重点放在了条约文本的设计细节上,其中不乏对条约的个案说明和对个别条款的应用规律研究。^②但截至目前,学界仍缺乏相对完整具体的分析框架。本文尝试建立条约约束力评价体系,包括五大维度的20个具体指标,运用文本分析及熵值法的计算方法对条约的约束效力进行量化分析,并以中国与东盟国家签署的海洋合作条约为应用对象检验该评价体系。

一 文献回顾

条约约束分析的研究方向集中在条约整体的约束分析和条款设计规则的约束分析两方面。

(一) 条约整体的约束分析

针对条约整体的约束分析主要在国际法和国际关系学科内展开。首先,国际法学科中的实证主义对国际条约的有效性进行了简单的二元区分,也有少数学者拒绝承认软法。卡尔·劳斯迪亚(Kal Raustiala)认为,合同(contract)和承诺(pledge)之间的最大差别在于使用法律和拒绝使用法律,一项国际条约只存在有效和无效两种情况,不存在中间地带,国家与非国家行为体(国际机构或社会组织)签署的条约以及不具备监督执行要素的条约均是无效的。^③但多数国际法学者承认软法概念的存在,认为国

^① Judith Goldstein, et al., "Introduction: Legalization and World Politics," *International Organization*, Vol.54, No.3, 2000, pp.385-399; Gregory Shaffer and Mark A. Pollack, "Hard and Soft Law," in Jeffrey L. Dunoff and Mark A. Pollack, eds., *Interdisciplinary Perspectives on International Law and International Relations the State of the Art*, New York: Cambridge University Press, 2013, pp.197-243; Kal Raustiala, "Form and Substance in International Agreements," *American Journal of International Law*, Vol.99, No.3, 2005, pp.581-614.

^② 随着冷战结束以及国际规则和机构的作用日益突出,国际法和政治学之间相互忽视的现象开始消退。1989年法律学者肯尼思·阿博特发表了一份宣言,呼吁对国际法进行跨学科研究,并鼓励法律学者借鉴最近的政治学学术成果。在接下来十多年研究的中,越来越多的法律学者开始对国际法的设计和细节进行研究,并借鉴国际关系理论以及定性/定量方法进行案例分析。2000年是一个理论转折点,《国际组织》第54卷和第55卷有两篇文章专门讨论了关于国际法的细节问题。参见 Judith Goldstein, et al., "Introduction: Legalization and World Politics," pp.385-399; Barbara Koremenos, Charles Lipson and Duncan Snidal, "The Rational Design of International Institutions," *International Organization*, Vol.55, No.4, 2001, pp.761-799.

^③ Kal Raustiala, "Form and Substance in International Agreements," pp.581-614.

家与非国家行为体之间的法律是软法。虽然软法没有正式的法律约束力,但属于为促使某种形式的说服而施加的一种依从约束力,是一种次优选择或未来的硬法。^①但上述观点均未对如何划分条约的“软硬程度”进行明确说明。

其次,在新制度主义影响下,近年来国际关系学科中出现了一系列国际条约效力判别的研究成果,其中理性选择理论更是条约有效性的核心理论支撑。肯尼思·阿博特(Kenneth W. Abbott)和邓肯·斯奈德尔(Duncan Snidal)认为,可以通过条约设计中所体现的义务程度(obligation)、精确度(precision)和委托程度(delegation)分析一份条约的有效性,并通过确定这三个要素程度大小所产生的成本—收益模式解释条约合法化的差异。在不同的条件(不确定性水平、时间范围等)下,硬法和软法会带来不同的成本—收益比率。阿博特和斯奈德尔将三要素分为高、中、低三个等级,认为只有义务程度高、精确度高、委托程度高以及义务程度高、精确度适中、委托程度高两类组合可以被视为硬法即强制性国际法,^②其余均应视为软法。除明确界定软法和硬法的概念外,阿博特和斯奈德尔还将每一种可能出现的组合进行举例说明,提出了首个较为完整的条约效力判定框架。^③芭芭拉·凯里迈诺斯(Barbara Koremenos)等基于制度主义学派对国际条约文本细节的评价,发展完善了理性设计框架,建立了国际法大陆研究框架(COIL),并将最初对国际组织有效性的五个维度(成员资格、议题范围、集中度、控制度和灵活性)分析发展成对国际条约有效性的研究。凯里迈诺斯等同样假定国家为理性的行动者,但基于博弈论的观点认为合作至少会使一方利益得以增加而另一方利益不受损害,违背协议则将给国家造成名誉损毁、利益被剥夺等不利影响。这种理性设计不讨论国家是否可能进行合作,而是讨论在确定合作的背景下国家面对不同合作问题时如何确定条约的规则。为探究条约的完整性与解决争端条款的关系,凯里迈诺斯等将条约的完整程度通过条约的复杂度(试图解决的问题非常简单、一般简单、一般复杂还是非常复杂)、精确度(条款内容的精确程度非常模糊、一般模糊、一般精确还是非常精确)、条约长度、是否包括附件以及是否引用其他国际法等指标进行量化说明。其通过计量分析得出结论:条约的完整性与要求对解决争端进行委托的条款呈负相关,且在标准误 $\rho=0.01$ 水平显著,而与规定仲裁或裁决解决争端的条款没

^① Gregory Shaffer and Mark A. Pollack, “Hard and Soft Law,” in Jeffrey L. Dunoff and Mark A. Pollack, eds., *Interdisciplinary Perspectives on International Law and International Relations the State of the Art*, pp.197-243.

^② 委托通常是处理不完全控制决策的最佳方法。阿博特和斯奈德尔认为,只要存在明确的责任义务要求和第三方委托,即便条约并非完全精确,这样的制度设计也已能够利用行政和司法机构解释和扩展其法律原则,所以义务程度高、精确度适中、委托程度高的组合同样应被视为硬法。

^③ Kenneth W. Abbott and Duncan Snidal, “Hard and Soft Law in International Governance,” *International Organization*, Vol.54, No.3, 2000, pp.421-456.

有显著关系。^①此外,银红武通过六语步的分析方法研究了条约文本约束。他统计了各语步出现的频率,认为国际条约语步三“宣告共识”的语言表达相对有力地佐证了国际法学界关于“条约的约束力来自缔约方协调意志”的结论,可作为条约有效性的证据之一。^②

(二) 条款设计规则的约束分析

近年来更多的学者将研究重点放在了条款设计规则的约束分析上。劳伦斯·赫尔弗(Laurence R. Helfer)针对条约的形式和实质,认为一份条约的设计是在灵活和烦琐的选择中进行权衡。灵活的条款虽然能够针对未来合作环境的变化随时调整以确保合作及时达成,但国家也可以援引这些灵活条款,在面临政治、经济和文化问题时规避害处(辩解),产生投机行为。过于烦琐的条款虽然会降低国家签署条约的意愿,但却能确保条约签署后保持强约束力。^③赫尔弗列举了灵活条款的种类,包括保留/声明条款、生效的要求、地区的限制、时间规定、条约修订的规则以及暂停和退出终止的规则。他认为灵活条款的介入有时无法达到公正的目标,反而将加剧国家间合作的不平衡性,有时甚至会偏离条约起草者的期望和学者的设想。^④基德·安德鲁(Kydd Andrew)认为,一份条约中对于成员资格要求的条款是作为一种信息和条约约束保证的手段。严格的成员资格标准不仅为选择良好的联盟伙伴提供了有用的信号,而且对监督成员国与第三方国家的合作也具有重大影响。^⑤凯里迈诺斯和艾利森·诺(Allison Nau)在对退出条款的“通知期(notice period)”和“等待期(waiting period)”进行定义和统计后发现,退出条款在人权条约中出现的频率最高:在随机样本的23项人权条约中有20项包含退出条款,占比87%;而经济条约、环境条约和安全条约包括退出条款的比例分别为57%、64%和44%。他们基于国际法大陆研究框架判断了在这些条约中是否存在执行问题和承诺问题,利用普通二乘回归(OLS)和probit分析发现:存

① Barbara Koremenos, Charles Lipson and Duncan Snidal, "The Rational Design of International Institutions," pp.761-799.

② 银红武:《人类基本价值的确信:国际条约约束力基础之语步结构实证研究》,载《时代法学》,2012年第3期,第109页;银红武:《国际条约约束力建构之法律体裁研究》,载《湖南财政经济学院学报》,2013年第3期,第10页。

③ Laurence R. Helfer, "Flexibility in International Agreements," in Jeffrey L. Dunoff and Mark A. Pollack, eds., *Interdisciplinary Perspectives on International Law and International Relations the State of the Art*, pp.175-196.

④ 在某些情况下,正式和非正式的灵活性机制之间存在预测不到的替代效应。例如,各国在《关贸总协定》中使用了第19条中的免责条款并将其作为变通工具。第19条并没有发挥预期作用的原因之一是国家更加频繁地采取了“灰色地带”措施,即一个关贸总协定成员说服另一个成员“自愿”限制出口或同意其他形式的管理性措施。

⑤ Kydd Andrew, "Reputation and Cooperation: Guzman on International Law," *International Theory*, Vol.1, No.2, 2009, pp.295-305.

在执行问题的条约的退出通知期比没有执行问题的条约更长,存在承诺问题的条约的退出等待期比没有执行问题的条约更长。^①冯寿波认为国际法中的国内位阶比照论具有一定意义,不同名称应附有不同的法律约束力,^②他根据《缔结条约程序法》认为在缔约/批准/审查权限方面,主体权力不同将会影响条约名称,如造法性条约常使用“公约”这个名称。^③陈兆源等则分析了中国签署的双边投资条约,发现中国在作为东道国时,自身经济形势发展越好,越容易接受高制度化水平的争端解决机制;中国作为母国时,经济预期越好的东道国越容易接受争端解决机制的委托行为。^④

综上,大部分对于条约约束力的定量研究均偏向于单项条款设计的因果分析,对于条约总体约束力的分析框架也过于笼统。虽然阿博特和斯奈德尔建立了三个维度(义务程度、精确度和委托程度)的条约有效性评估框架,但更多也是以举例方式进行说明和定义,未能给出明确的定量评价方法。为此,本文试图构建起一个相对完整具体的分析框架,对国际条约约束力进行量化分析。

二 国际条约约束力的计算

国际条约约束力框架的建立主要基于条约文本内容的判断,本部分将阐述建立框架所需的理论,并在此基础上对如何确定框架内指标进行逐一说明,最终给出条约约束力的具体计算方法。

(一) 理论基础

阿博特和斯奈德尔的分析框架具有开创性意义但有待进一步具体化,相比之下,凯里迈诺斯在这一框架基础上发展出的理性设计框架(国际法大陆研究框架)是一套相对完整的条约文本统计分析体系,该框架利用对条约文本信息的500多个问答评价和信息记录,在数据统计基础上进行条约设计规则的因果验证。本文选择凯里迈诺斯的理性设计框架作为条约约束力评价体系的理论基础,并通过筛选和补充操作性更强

^① Barbara Koremenos and Allison Nau, "Exit, No Exit," *Duke Journal of Comparative & International Law*, Vol.21, No.8, 2010, pp.81-119.

^② “法律位阶”是指每一部规范性法律文本在法律体系中的纵向等级。下位阶的法律必须服从上位阶的法律,所有法律必须服从最高位阶的法。国内法位阶比照论认为,尽管在一项国际条约中,不同国家针对某一国家具有同等法律效力,但从国内层面来看,条约缔结和解释权力实际上由国内法规定。因此,不同名称的条约在国内的不同位阶可能影响条约国内解释主体的确定。

^③ 冯寿波:《论我国条约解释主体制度的完善——以修订我国〈缔结条约程序法〉为视角》,载《政治与法律》,2014年第9期,第10—23页。

^④ 陈兆源、田野、韩冬临:《双边投资条约中争端解决机制的形式选择——基于1982—2013年中国签订双边投资协定的定量研究》,载《世界经济与政治》,2015年第3期,第122—148页。

的评价指标,尝试建立一套完整的条约约束力评价体系。

基于凯里迈诺斯的理性设计框架,本文从成员资格(membership)、议题范围(scope)、集中度(centralization)、控制度(control)和灵活度(flexibility)五个维度对条约约束力进行了考察。

一是成员资格。成员资格是指条约签署成员国的特征,判断条约是否具有排他性和限制性。例如,国际公约从最初的条约形成就具有包容性,几乎任何国家都可以参与签署,无论其经济发展程度如何、区域位置在哪里,而一些区域性合作条约如《南印度洋渔业协定》对签署成员的地理位置有一定要求。另外,还有一些条约因其合作内容的特殊性而对条约的签署国有一定的经济和技术要求。一份条约的签署往往要求签署国承担相对应的权利与义务,签署成员的数量和能力都会对条约的约束力产生一定影响。

二是议题范围。在众多的国际条约中,既有大到作为某一领域的共识性公约(任何相关领域的条约都可以参照或者引用这样的公约),又有小到仅仅讨论一个具体细则的条约。一份条约的主题可能涉及很多领域,也可能是为建立某个合作组织而形成的。例如,《南极条约》涵盖了一系列科学、经济和政治问题,多领域的合作使得国家必须全面配合,就条约各个方面的要求进行相应调整和谈判,这在无形中增加了条约执行的成本。但一些早期的环境协议仅涉及明确排放界定之类的问题,如温室气体的定义和具体的排放限制,签署成员可以就某一项需要调整的问题进行具体讨论,这有助于增加条约执行的效率。可见条约所涉及问题的多寡是一份条约实施的关键影响因素,且不同主题条约存在较大的设计差异。

三是集中度。主权国家通常拒绝任何形式的国际集权,以维护自己国家行动的独立性和控制力。集中度主要包括信息搜集、降低协商成本和加强执法三方面。其中,信息搜集主要指在条约执行过程中需要通过了解合作国家的具体信息以保证条约的持续有效执行。如在1968年《国际咖啡公约》中,为控制咖啡市场供给和维持咖啡价格,巴西、哥伦比亚和非洲国家需要通过搜集对方的海关数据以保证各方在出口中遵守了条约。又如在有关环境保护的条约中,签署国也需要搜集各国的排放情况确保减排目标的达成。降低协商成本是指在对某一部分条约进行修改或者通过特定议题时,选择集中会议方式会比双边协商更加节约谈判和执行成本。加强执法则是行为监督和对违背条约的惩罚,尤其是在争端解决方面存在一个可靠的执法者将会对履约行为形成有力约束。

四是控制度。控制由一系列因素决定,主要包括选举重要官员的规则和机构的融资方式。控制度与集中度的侧重点不同:集中度是指执行条约时所体现的执行方式,控制度则是指决策阶段所选择的权力分配方式。这种分配方式是完全平等的还是具

有大国权力政治色彩、是根据某方面能力水平分配还是基于权利义务均等分配,都对条约本身的约束力具有一定影响。如在针对国际性条约的决策投票规则中,所有成员都有平等的投票权;而在一些经济合作类型的条约中,对合作提供更多资金帮助的成员会被分配更多的加权票数。如果少数签署国就可以决定条约中的决策,则意味着它们的选票本身就具有特殊的分量,合作的约束将会因少数签署国的行为减少,使得其他签署国的合作利益存在不确定性。

五是灵活度。随着国际环境的不断变化,条约中的许多条款很可能将不再适用。针对条约设计的研究,争议最大的地方就是关于条约灵活度的设计。同其他学者认为灵活度有助于国家遵守条约相比,凯里迈诺斯指出更加灵活的条款在追责的过程中往往难以执行。国家可能会通过各种各样的灵活条款(免责条款和保留条款等)对自己的违约行为进行辩解,这降低了条约对于国家的约束力。

凯里迈诺斯将研究重点放在条款与合作问题的相互影响上,即在不同合作背景和出现不同合作问题时条款是如何被选择的。在对条约 500 多个特征的评价中,很多是研究人员利用自己的专业知识进行打分,这具有一定的主观色彩。与凯里迈诺斯的理性设计框架不同,本文将研究重点放在条约本身,既不考虑国际合作是否会发生(签署前),也不考虑国际合作是否会被违约(签署后),仅仅通过条约文本内容对条约的约束强度进行评价。这种用相同的标准来判断的方法有助于分析不同国家与不同类型的条约条款设计的差别,也可以通过研究同类型条约设计的变化规律探索国家间合作关系的微妙变化。

(二) 指标框架的设计

经过对有关条约问题的资料整合,本文最终筛选出 20 个较为客观的问题作为五大维度的具体衡量指标,并对条约条款的约束力进行区分和赋值,其中最小值为 0,差值为 1,按照约束越强分数越高的赋值标准进行评价(见表 1)。下文将对具体的赋值标准进行讨论。

表 1 条约约束力指标框架

维度	项目	条款	内容	赋值
成员资格	成员条件	条约类型	双边	2
			封闭多边	1
			开放多边	0
		成员准入条件	严格	2
			宽松	1
			无	0

续表 1

维度	项目	条款	内容	赋值
议题范围	主题	主题细致程度	具体情况	2
			具体主题	1
			笼统	0
集中度	信息	信息搜集	机构搜集	2
			自行上报	1
			没有要求	0
		信息审核	机构审核	1
			没有要求	0
		大会	审查条款	有
	无			0
	执行	术语条款	定义必须为	2
			定义为	1
			无	0
		行为规范	具体	2
			模糊	1
			建议	0
		行为监督	有	1
			无	0
		最高惩罚方式	被开除	3
			被惩罚/权利削弱	2
			应当惩罚(建议)	1
			没有要求	0
		国内法规调整	有	1
			无	0
		争端解决方式	裁决	3
			仲裁	2
			调解	1
非正式			0	
争端解决途径	外部机构	2		
	内部机构	1		
	无	0		

续表 1

维度	项目	条款	内容	赋值
控制度	投票	票额分配	非权重	2
			权重	1
			特殊	0
		最低决议方式	一致通过	3
			绝对多数	2
			简单多数	1
			特殊	0
灵活度	灵活条款	免责条款	无	1
			有	0
		退出条件	无/禁止退出	2
			附加条件	1
			无条件	0
		退出通知期	有	1
			无	0
		退出等待期	有	1
			无	0
		保留条款	禁止保留	2
			附加条件	1
			无条件/无	0

资料来源:笔者自制。

各项指标赋值依据及标准如下:

1. 成员资格

成员资格主要包括条约类型和成员准入条件两个指标。

一是条约类型。无论是条款修改还是责任追究,签署国的数量越多越可能增大条约执行的不确定性,导致条约的约束力下降。因此双边条约通常比多边条约具有更高的约束力,^①本文因而赋予双边条约最高值。在对多边条约的分类中,本文采纳联合国条约汇编(UNTS)数据库的处理方式将其分为封闭性多边条约和开放性多边条约两类:封闭性多边条约在签订之后不会再有新的国家加入,并且一般是区域性的,本文赋

^① Barbara Koremenos, Charles Lipson and Duncan Snidal, "The Rational Design of International Institutions," pp.761-799; Thomas Rixen, "Bilateralism or Multilateralism? The Political Economy of Avoiding International Double Taxation," *European Journal of International Relations*, Vol.16, No.4, 2010, pp.589-641.

予其比开放性多边条约更高的分值。综上,双边条约赋值为2,封闭性多边条约赋值为1,开放性多边条约赋值为0。

二是成员准入条件。为保证条约内容的顺利实施,一些多边条约会对签署国的准入条件提出一定要求。如《建立印度洋金枪鱼委员会协定》要求后续申请加入该委员会的国家应当是联合国粮农组织的成员国,且经2/3以上的原有委员会成员国投票同意才可以加入。1974年联合国《班轮公会行动守则公约》中也明确规定,有意愿申请加入的国家及公司应提供相关能力的证明保证其特许吨位情况符合本条款标准。而对于那些只要申请就可以签署的条约,其执行效果可能因国家能力的不足或国内政治环境的混乱而大打折扣,后期也无法承担相应的违约费用。此外,本框架对不同的准入要求也做了进一步划分。在上述两个例子中,显然前者对条约执行不能提供任何客观保障,而后者对于能力的要求则更加严谨。尽管在第一种情况中,联合国粮农组织成员国的申请也需要一定的能力证明,但在《建立印度洋金枪鱼委员会协定》中,并未提及关于联合国粮农组织成员国的申请条件,本框架仅对条约文本中所包含的具体文字条件进行分析,并不考虑暗含的内容和规则。与多边条约相比,双边条约是两个国家签署的具有针对性合作的条约,对象国只有一个,其目的和签署意愿明显,针对性很强,因此将双边条约的成员准入条件定义为“严格的准入条件”。综上,严格的准入条件(有能力要求和双边条约)赋值为2,宽松的准入条件(有条件但无能力要求)赋值为1,无要求(或未提及)赋值为0。

2. 议题范围

议题范围主要关注主题的细致程度。凯里迈诺斯等发现,主题的异质性随着成员国数量的增加而上升,随着合作中出现执行问题的严重程度加大而上升。^① 条约内容中包括的主题越多样,越能反映出条约在执行问题方面面临问题的严重程度。为解决不同偏好的国家在合作中的收益不平衡问题,开展合作的国家往往通过增加合作内容进行利益交换。因此,一份条约所包含的主题内容越多越复杂,则越能说明条约签署国的合作问题越多;而合作问题越多,合作的不确定性就越大,条约的约束能力也就越弱。反之,一份条约如果只为解决一个问题,签署国的异质性就越低(为解决相同问题而签署),目标也更加明确,约束力也就越强。而那些涉及国家间全面合作的国际条约,条款内容通常包括多个子领域问题,其执行过程相较于单一子领域就更为复杂。此外,那些为建立组织机构或委员会而形成的条约,其约束力就更加低下。综上,条约

^① Barbara Koremenos, Charles Lipson and Duncan Snidal, "Rational Design: Looking Back to Move Forward," *International Organization*, Vol.55, No.4, 2001, pp.1051-1082.

目的为解决单一问题的赋值为 2,为解决单一子领域问题的赋值为 1,为解决多个子领域问题/开展全面合作/建立组织机构/成立委员会的赋值为 0。

3.集中度

本文沿袭凯里迈诺斯的做法将集中度分为信息、大会和执行三方面。信息方面包括信息搜集和信息审核,大会方面包括审查条款,执行方面包括术语条款、行为规范、行为监督、最高惩罚方式、国内法规调整、争端解决方式和争端解决途径,总共有 10 个具体指标。

一是信息搜集。在双边和多边条约中,对条约执行过程中信息搜集的要求主要包括两种方式:一是指定机构对签署国进行信息搜集,二是签署国自行上报信息。其中,指定机构包括指定外部机构或内部机构,其区别在于外部机构是委托的第三方机构,内部机构是由签署国为某种目的而成立的委员会或组织等。本文在信息搜集方面并未对外部机构和内部机构进行约束强度的区分,原因在于大部分条约没有对信息搜集做出要求,而要求机构进行信息搜集的条约就更少,因此本文将内部机构和外部机构视为同一程度的约束。同时在一一般情况下,国家自行上报信息会因为签署国之间的数据规则不统一而出现误差,有时甚至会出现国家瞒报谎报的情况,因此体现的执行约束更低。综上,机构统一搜集赋值为 2,国家自行上报赋值为 1,未提及赋值为 0。

二是信息审核。前文提到,签署国自行上报存在瞒报谎报的情况,但是国家自行上报的数据如果经过严格监督和审查则可以避免这种情况的发生。因此,针对搜集到的信息进行审查将具有更强的执行约束,也更能保证所搜集信息的质量。本文将有信息审核制度赋值为 1,未提及或没有信息审核制度赋值为 0。

三是审查条款。审查条款是指条约中是否规定了成员国需在一定时间内召开大会,以便监察现有条约的执行力度及对现有条约在执行过程中出现的问题进行修正。这种行为会增强对条约的执行约束力。例如,《养护大西洋金枪鱼国际公约》第三条第 4 款要求委员会每两年举行一次会议。^① 本文将有审查条款赋值为 1,无审查条款赋值为 0。

四是术语条款。尽管对于条约术语的解释理论是必需的,但现实中仍存在一些条款并没有对条约中出现的词语做出定义解释。在国际合作中,国家违反承诺既可能是由于缺乏遵守的能力,也可能只是因为对语言理解的不同而对条款的定义产生

^① United Nation Treaty Collection, "International Convention for the Conservation of Atlantic Tunas," https://treaties.un.org/Pages/showDetails.aspx?objid=0800000280124e44&clang=_en, 访问时间:2021 年 10 月 6 日。

误解。^① 在关于世界贸易组织(WTO)的争端中,相当一部分是因为语义歧义所产生的合作纠纷。如在2003年安提瓜和巴布达诉美国网络赌博服务案中,在安提瓜和巴布达与美国签署的条约中由于未对“体育(sport)”一词进行定义导致双方产生不同的语义理解而发生争端。可见,有无术语解释会对条约执行产生重要影响,因此有必要增加此项评价指标。

除未做出规定所导致的语义分歧外,还有一些条约故意模糊术语解释,让部分国家可以在那些与自身文化、制度相冲突的条款中做出解释,从而放宽原则。《美洲人权公约》由美洲国家组织制定,其中第4.1条对生命权的规定中要求“人人有权使其生命得到尊重,每个人都有权让自己的生命受到尊重。这项权利从受孕之日起,一般应受法律保护。任何人的生命不得任意剥夺”。^② 但因为美洲国家组织的23个成员国中有12国的法律允许在某些情况下堕胎,所以在最后条款中加入了“一般情况下(in general)”的限定条件。由此可见,术语定义的强弱也对其约束力大小具有重要影响。冯寿波的研究发现,在法律英语中使用频率最高的情态动词依次是shall、may、must和should。在WTO案件的裁决报告中,当某些案例涉及定义内包含shall和should的概念辨析时,明确规定了二者不同的准司法解释。如在欧洲共同体关于肉和肉制品(荷尔蒙)案的专家组报告中明确提及,shall与should的最重要区别为是否具有法律约束力。在法律语言中,shall表示命令、义务和职责等,其作用相当于must、be required to、be to (do)与have to,在条约中文版中一般会将shall译为须、应、应当或应该等。^③ 如果没有对术语定义的强烈程度进行区分,根据条约解释的习惯规则,将以“遇有疑义从轻解释(in dubio mitius)”为标准,^④ 从而降低条约的规范程度。有理由认为,当术语定义出现shall、must、be required to、be to (do)、have to、须、应或应当等词汇时赋值为2,没有出现上述词汇但有对术语进行定义时赋值为1,没有提及定义时赋值为0。

五是行为规范。对行为规范的内容考察也是衡量条约精确性的重要组成部分。若条约对签署国之间的权利与义务分配更加明确,对责任的追究也就更加清晰。在对行为规范的描述中,条约中出现“完善建设”“共同努力”“争取打造”等较为空泛的词

① Barbara Koremenos, “Institutionalism and International Law,” in Jeffrey L. Dunoff and Mark A. Pollack, eds., *Interdisciplinary Perspectives on International Law and International Relations the State of the Art*, p.71.

② United Nation Treaty Collection, “American Convention on Human Rights ‘Pact of San José, Costa Rica’,” https://treaties.un.org/Pages/showDetails.aspx?objid=08000002800f10e1&clang=_en, 访问时间:2021年10月6日。

③ 冯寿波:《〈WTO条约〉实证研究中“shall”和“should”词义研究》,载《西南政法大学学报》,2013年第5期,第33页。

④ 即条约解释时如遇用词含糊不清的情形,则应优先选取对承担义务的当事方影响较轻、限制较少的含义。

汇时,即代表未针对具体问题提出合规建议,在没有任何定量指标的要求下,条约的完成程度无法测量,将导致条约目标被无限拖延甚至陷入停滞;还有一些条约的条款中,并未提出合作国家的行动目标,仅仅作为规劝性公约的存在,对国家行为的规范也只是止于建议,约束就更加低下。综上,当条约中具有精确的行为规范(包括具体数字的定量要求)时赋值为2,只有模糊的行为规范(不包括具体数字)赋值为1,建议类规范赋值为0。

六是行为监督。在信息不对称的情况下,国家汇报的行为与其实际执行的行为可能存在偏差。当统一的机构主体对国家行为进行监督时,国家必须让渡部分主权以保证对条约的遵守程度。此外,有监督要求的条约也能够针对出现争端的情况进行公正的评价。例如,《关于执行1982年12月10日〈联合国海洋法公约〉第十一部分的条约(附附件)》第三部分第5款(c)针对工作进度就有对“监测勘探工作计划的遵守情况”的要求。^① 本文将有行为监督要求赋值为1,无行为监督要求赋值为0。

七是最高惩罚方式。惩罚是针对行为监督后采取的措施,设立监督机制是为了检查条约是否被很好地遵守,惩罚机制则是针对监督结果的执行。当一国违反了条约内容,其他条约签署国及其设立的机构有权追偿。一份条约会针对不同的违约行为采取不同的惩罚措施,最高程度的惩罚措施可以展示条约本身的约束程度。因此,可以将最高惩罚方式作为最终约束程度的判断依据。根据制度主义理论,国家违约会对自身造成一定的损失,因此违约的成本越高,国家越不愿意违约,条约的约束力就越强。而开除成员资格会让违约国直接丧失其在条约内的所有权利,无疑是最严厉的惩罚措施。综上,当最高惩罚方式为直接除籍或终止合作时赋值为3,部分权利丧失或受到其他惩罚时赋值为2,建议进行惩罚(并未采取强制性措施)时赋值为1,未提及惩罚时赋值为0。

八是国内法规调整。在对国际法约束力的研究中,也有相当一部分学者从国际法与国内法的关系层面进行讨论。^② 国际条约的签署可能会对国内法规产生两类影响:一是一国根据不断签署的国际条约,按照国际标准倒逼国内法规的改革;二是在未对国内法规做出调整要求时,面对国际条约内容与国内法规的冲突,不同国家采取不同的应

^① United Nation Treaty Collection, "Agreement Relating to the Implementation of Part XI of the United Nations Convention on the Law of the Sea of 10 December 1982," https://treaties.un.org/Pages/showDetails.aspx?objid=0800000280045372&clang=_en, 访问时间:2021年12月6日。

^② Judith Goldstein, et al., "Introduction: Legalization and World Politics," pp.385-399;李浩培:《条约法概论》,法律出版社1987年版,第30页;赵建文:《国际条约在中国法律体系中的地位》,载《法学研究》,2010年第6期,第190页;曲亚因、韩立新:《论国际条约与国内法的关系》,载《东北师大学报(哲学社会科学版)》,2013年第6期,第254—256页。

对方法(位阶比照方法的不同),这无疑会破坏条约的执行。^①虽然《维也纳条约法公约》明确规定国家不得以国内法规为由而拒绝履行条约义务,但现实中此类纠纷往往难以追责。^②综上,当条约中明确提及国际条约与国内法规的关系时赋值为1,未提及时赋值为0。

九是争端解决方式。当条约签署国在执行条约的过程中出现冲突时,一般会按照条约中规定的争端解决方式进行调解。^③在国际合作中,条约规定的争端解决方式主要分为四种,包括外交与友好协商(diplomacy, friendly negotiation)在内的非正式手段、调解(mediation)、仲裁(arbitration)以及裁决(adjudication)。这四种争端解决方式的约束力是逐渐增强的。调解是比非正式手段稍微严格一些的方式,调解人通常只传递争论者的立场,但他们可以在当事人信任调解人的情况下预设不同的立场。仲裁是相对调解而言的一种更有效的争端解决方式,因为仲裁是由争端各方选定的第三方来解决争端,通常会规定如何选择仲裁员以及其判定是结论性的(即对当事人有约束力)还是简单的建议。而约束力最强的裁决则是直接对结果进行判决并强制实施。^④一项条约中对于争端解决方式的规定往往不止一种,有时是不同问题采取不同的争端解决方式,而有时是同一个问题搭配两种争端解决方式。例如,英国和埃及之间的一项投资条约中规定“两国如果不能以外交手段解决争端,则必须将争端提交仲裁机构”。^⑤据此,本文以最严格的争端解决方式为评价标准。当最严格的争端解决方式裁决赋值为3,仲裁赋值为2,调解赋值为1,非正式手段赋值为0。

十是争端解决途径。在三种正式的争端解决方式中,条约有时会规定执行主体将

① 不同国家对于国际法与国内法的关系判断有所不同,主要分为一元论和二元论两类学说。一元论学说中又存在两种观点:一种认为国内法效力完全优于国际法,以黑格尔为代表,代表国家是希腊;另一种认为国际法完全优于国内法,以康德为代表,代表国家是阿根廷。二元论学说认为国际法和国内法隶属不同的法律体系,宪法、一般法和国际法具有不同的优先程度:一种观点是国际法低于宪法,与一般法同等,代表国家为美国、墨西哥;另一种观点认为国际法低于宪法、高于一般法,这也是最常见的处理方法,代表国家为法国、爱沙尼亚、俄罗斯和中国等。参见曲亚囡、韩立新:《论国际条约与国内法的关系》,载《东北师大学报(哲学社会科学版)》,2013年第6期,第254—256页。

② 1969年《维也纳条约法公约》第26条规定“凡有效之条约对其当事国有拘束力,必须由各该国善意履行之”;第27条规定“一当事国不得援引其国内法规定为理由而不履行条约。此项规则不妨碍第四十六条”。参见United Nation Treaty Collection, “Vienna Convention on the Law of Treaties,” https://treaties.un.org/Pages/showDetails.aspx?objid=080000028003902f&clang=_en, 访问时间:2021年10月14日。

③ Judith Goldstein, et al., “Introduction: Legalization and World Politics,” pp.385-399.

④ Barbara Koremenos, “The Design of Dispute Settlement Procedures in International Agreements,” in Jeffrey L. Dunoff and Mark A. Pollack, eds., *Interdisciplinary Perspectives on International Law and International Relations the State of the Art*, pp.371-374.

⑤ United Nation Treaty Collection, “Agreement for the Promotion and Protection of Investments,” https://treaties.un.org/Pages/showDetails.aspx?objid=08000002800ff81&clang=_en, 访问时间:2021年10月6日。

条约执行委托给内部机构或第三方外部机构。一般而言,在解决争端时,委托给外部机构可以提高条约的效率和公正性。^① 据此,将争端解决委托给外部机构赋值为2,委托给内部机构或国家自行解决时赋值为1,未提及争端解决条款赋值为0。^②

4. 控制度

控制度考察的是条约在执行过程中通过某项决定时所采取的投票规则,这也是对国际关系现实主义理论所强调的国家权力的一种侧面体现,主要有两个衡量指标。

一是票额分配。投票规则涉及制度内部的权力关系,越平等的票额分配越能削弱大国在条约中的控制力,能够通过增加国家间的相互牵制以提高条约的约束力,反之亦然。如《国际水文组织公约》第11条第4款中规定“各成员国政府应享有由其船队吨位所定比例的票额分配”,^③这种以实力确定配额的分配方式在贸易和投资等低级政治领域的合作条约中最为常见。票额分配也会以组合的方式出现:在每国n票的基础上,将多余的票额分配给贡献最多的国家,这是一种非权重加权重的投票规则。因为票额的分配有多种组合方式,因此,本文按照组合中最能产生权力差异的票额分配方式来考察。据此,对于双边条约,当条约中出现“平等(equality)”等表述国家合作地位相等的词汇时视为非权重原则,此时赋值为2;当条约出现按权重分配或未提及权重时分值为1;当条约为无偿帮助条约时,是一种权利义务不对等的条约,施助者拥有特殊主动权,赋值为0。对于多边条约,当仅存在一国一票规则时赋值为2,包括权重票额但无特权票额时赋值为1,包括特权票额或未做出说明时赋值为0。

二是决议方式。投票决议的权力分配还包括决议规则,主要有一致通过(unanimity)、绝对多数(super majority)和简单多数(simple majority)三种。在一致通过的决议方式下,通过一项决议需要全部成员国的同意,尽管达成一致的过程比较漫长,但一旦敲定结果便不会有太大分歧。而在多数通过决议方式下,部分大国可以向摇摆不定的小国施加压力以求达成对自己有利的决议结果。决议规则对国家行为控制起到了一定的作用,所需通过决议的票数要求占比越大,在条约后期产生的分歧就越小,条

^① Barbara Koremenos, “When, What, and Why Do States Choose to Delegate?” *Law and Contemporary Problems*, Vol.71, No.1, 2008, pp.151-192; Curtis A. Bradley and Judith Kelley, “The Concept of International Delegation,” *Law & Contempt*, Vol.7, No.1, 2008, p.1.

^② 设置0值是为了区分条款中明确规定争端通过友好协商解决和未对争端解决机制做出规定的条约,当条款中规定争端解决机制为非正式时,虽然在争端解决方式这一评价指标未得分,但在争端解决途径中可以赋值为1,完全未提及的条约则赋值为0。

^③ United Nation Treaty Collection, “Convention on the International Hydrographic Organization (with annexed General Regulations and Financial Regulations),” https://treaties.un.org/Pages/showDetails.aspx?objid=080000028011f30b&clang=_en, 访问时间:2021年10月12日。

约的约束力也就越大。^① 在一份条约中往往针对不同的问题存在不同的决议方式,本文以最宽松的规则作为评分依据。在双边条约中,要求必须两国一致通过才能达成条约,有任何一方拒绝条款,条约都将失效。因此,所有的双边条约在这一项均赋值为3。对于多边条约,当最低决议方式为一通过时赋值为3,为绝对多数通过时(大于或等于2/3)时赋值为2,为简单多数通过时(小于2/3大于1/2)赋值为1,部分国家享有特权或未提及决议方式时赋值为0。

5. 灵活度

一份条约通常是在适应性(更加灵活)和保障性(更加稳固)中做出选择。当出现灵活条款的时候,条约的约束力也随之下降。免责条款和退出条款是灵活度的两种主要表达形式。其中,退出条款是在国家偏好改变时做出的灵活性选择,免责条款则是国家偏好不变、国内约束改变(如战乱、政变)时所能引用的灵活条款。此外,本框架增加了对保留条款的考察:通常情况下,一国会对一份条约中无法接受的条款进行保留,以求摒除或更改条约规定对声明人的法律效果,这同样是一种灵活条款。^② 本文根据三种条款,设定五个具体的衡量指标。

一是免责条款。免责条款一般是指由于受到国内的政治经济打击而“暂时”解除义务的条款,但为防止滥用,条约为免责条款增加了越来越严格的条件,如通过补偿或者信息披露以证明适用情况确实发生。在条约文本中,免责条款有时会以标题的形式进行具体说明,有时也会以“由于不可抗力(force majeure)”“对……没有责任(no liability for)”“免除义务(relieve obligation)”“除……情况外(except)”等表达方式出现。值得注意的是,要区分免责条款和例外条件(exception to compliance)。如《国际油污损害民事责任公约》第三条第2款对于免责部分的说明为“可以免除义务如果船主能证明:a.起因于战争、敌对行动、内战、暴动或具有特殊、不可避免和不可抗拒性质的自然现象”,对于强制执行的说明是进行了例外条件的说明,即“任何国家……不再需要进行复查……除非:a.以欺诈手段获得”。^③ 可见,例外条款是指遵守条约规定的规则在哪些情况下可以不予考虑、是永久性的,免责条款则是临时性的。无论如何,如前文所述,免责条款作为一种灵活条款都会导致条约约束降低,因为存在被强行引用以逃避条约义务的可能。据此,当条约中无免责条款时赋值为1,有免责条款时赋值为0。

① Barbara Koremenos, “When, What, and Why Do States Choose to Delegate?” pp.165-166.

② United Nation Treaty Collection, “Vienna Convention on the Law of Treaties,” https://treaties.un.org/Pages/showDetails.aspx?objid=080000028003902f&clang=_en, 访问时间:2021年10月14日。

③ United Nation Treaty Collection, “International Convention on Civil Liability for Oil Pollution Damage,” https://treaties.un.org/Pages/showDetails.aspx?objid=08000002801083db&clang=_en, 访问时间:2021年10月15日。

二是退出条件。对退出条款的评价主要考察退出条件、退出通知期和退出等待期三个指标。因为一国退出条约会影响或损害其他签署国的利益,对于一项双边条约来说,退出更是意味着两国合作的终止。通常而言,一份条约的退出条件主要分为六种类型:(1)允许随时退出;(2)允许在固定年限后退出;(3)允许在固定年限内退出;(4)允许在某种情况下退出;(5)完成条约内容后自动退出;(6)不允许退出。本文的框架中将这六种要求按照有无退出条件分为三大类,即任意退出、附条件退出和不允许退出。此外还有一种情况是没有对条约的退出条件做出说明,《维也纳条约法公约》第56条第1项规定,“条约如无终止退出条款,且未对退出做出规定的,适用不能退出”。因此,如果条约没有对退出做出说明,视为不允许退出。本文在条约不允许退出或没有对退出情况做出说明时赋值为2,有条件退出时赋值为1,任意退出时赋值为0。

三是退出通知期。退出通知期是指已签署条约的成员在想要退出条约时,发出退出通知与退出生效之间的一段时间。一般退出通知期越长,给条约中其他签署国留出的调整时间就越长。在退出通知期内,其他签署国可以在想要退出的成员退出前调整自己的政策,从而不至于在其退出后即刻蒙受利益损失。退出条件越多,退出成本越大,越能阻止签署国不负责任的退出行为,条约约束力也就越强。本文将有退出通知期赋值为1,无退出通知期赋值为0。

四是退出等待期。退出等待期是指退出条约的签署国能够完全摆脱条约义务的时间。^①例如,在一般的投资条约中,一国在退出条约后对其他未退出的签署国仍具有一段优惠保护期。在这段时间内,退出国家仍不得不对未退出条约的签署国提供投资优惠。这也是对条约签署国的一种权利保护机制,提升了条约对退出国家的约束力。据此,当条约有退出等待期时赋值为1,无退出等待期时赋值为0。^②

五是保留条款。保留条款的类型包括禁止保留、固定期限内做出保留、只允许对某一条款做出有条件保留和任意保留。其中如果未对条约的保留做出规范,国家可以根据《维也纳条约法公约》第19条做出保留,并口头或书面告知相关国家。本文将对条款禁止保留赋值为2,有条件保留赋值为1,可随意保留(未做说明)赋值为0。

6. 其他因素

本文除选取上述20个指标外,对条约有效期、条约名称和条约委托事项的区分也进行了考虑,但最终并未将其纳入框架指标,这主要基于三点原因。

^① Barbara Koremenos and Allison Nau, “Exit, No Exit,” pp.81-119.

^② 有些条约会要求国家在条约生效满一定年限时才可以提出退出申请,这在一定程度上延长了退出等待期。这种情况既属于退出条件中的“有条件退出”,也属于存在退出等待期,在赋值时都会予以考虑。

一是对条约有效期的排除。对于灵活条款,还存在一个重要的设计因素即条约的有效期限。条约的有效期分为有限和无限两类,通常以年为单位。一般情况下,经济类条约的有效期较短。因为经济指标一般以年为单位衡量,条约中的合作条件会根据每年的经济状况进行不同调整。而立法定性公约的有效期一般为无限期,有效期间会通过召开审查大会来修改不符合现状的条款,从而保证条约的约束力。在关于条约的有效期与其他条款的影响研究中发现,条约的有效期与其他灵活条款高度负相关,^①且退出条款已经从侧面反映出有效期的长短,因此不必纳入该框架重复衡量。此外,对于有效期限的时间划分目前并不能通过理论或实证证据说明其对于条约约束力的影响。综上,本文的指标框架中不包括对于条约有效期的评价。

二是对条约名称区分的排除。在最初的框架设计过程中,本文考虑过针对条约的不同名称进行约束力上的区分。本文所提到的国际条约均是广义上的条约,指任何主体之间缔约的法律文书。但在国际条约的分类中,确实存在对于名称的不同应用,包括公约、条约(狭义)、临时条约、专约、换文和决议等。有学者认为,使用不同名称的条约往往与该条约的内容与批准程序等存在一定相关性,在约束力上存在差异。^②但无论是在国际法经典教材还是在《维也纳条约法公约》对于条约名称的解释中,条约名称的不同都被认为不具有重大法律意义。^③在缺乏法律依据及实证证明的情况下,本文的指标框架未对条约名称进行约束力区分。

三是内部委托和外部委托问题。在本文的指标框架中,有关委托事项的评价只涉及信息搜集和争端解决两部分。凯里迈诺斯通过统计条约样本的条约委托内容发现,针对22个可委托的事项,委托应用最广泛的事项就是争端解决和信息方面的要求(包括信息搜集、信息审核、信息扩散和信息数据分析),因此本文并未对条约可能包含的全部委托事项进行考察。而对委托方式选择的统计发现,争端解决、信息搜集和信息扩散在外部委托和内部委托的选择上分布相对均匀,因此能够很好地区分条约之间的约束力。^④此外,本文排除了对信息扩散和信息数据分析的委托考察,是因为在实际形成的低级政治领域条约中几乎不包含这两项要求。

① Barbara Koremenos, "International Institutions as Solutions to Underlying Games of Cooperation," IBEI-Working Papers, B-21.147-2006, Barcelona, November 2009.

② 小约翰·金·甘布尔著,卢莹辉译:《多边条约各种名称的不同意义》,载《环球法律评论》,1983年第1期,第67—69页;冯寿波:《论我国条约解释主体制度的完善——以修订我国〈缔结条约程序法〉为视角》,载《政治与法律》,2014年第9期,第10页。

③ 《维也纳条约法公约》第二条第1款甲项规定,条约可采取其他特定名称,但并未明确不同条约名称在法律效力上的差别。

④ Barbara Koremenos, "When, What, and Why Do States Choose to Delegate?" pp.151-192.

(三) 评价方法

针对文本量化方法,本文将采取全变量观测法(AVFEC),即对一份条约的所有变量进行观测,而非按照一个问题对于全部条约逐一分。此种方法的优势在于:对于整份条约的阅读可以帮助评分员更多地了解条约的前后关系和细节,从而能够对一份条约做出更加准确的评分。尽管问题已经阐述得非常具体,但在条约约束力评价指标中仍有个别问题在人工打分时可能会出现偏差,在对文本评价的实践中通常建议使用多位可以比较的评分员。所以,在经过对评分体系的统一培训过后,每份条约都由至少两位评分员独立评价完成。在意见出现分歧时,评分员将对评价结果进行审查并协商解决。对于文本分析的准确性,通常需要将分歧控制在问题的2%—15%。

(四) 熵值法计算权重

在利用框架对条约进行基础评价后得到的分值并不能作为判断样本内条约约束力高低的最终值。学界对于条约约束不同维度的权重分配一直存在分歧,与其从理论上讨论如何赋予几个主要维度的具体权重,不如直接从条约文本的约束分值差异着手。在已经形成的某类条约文本中,可能存在某种约定俗成的条约设计规则,这些规则几乎存在于此类型条约的每一份条约中。在对条约进行评价时,条约约束力的差异性也会被那些约定俗成的规则影响而造成样本内的约束分值集中在某一区域,从而丧失对比研究的真正价值。因此,将权重赋予那些差异性更高的维度在条约约束力的对比研究中将更具说服力。熵值法是各评价体系中使用范围较广的评分方法,其根据特定样本数据的差异性对各评价指标的权重进行重新分配,以此得到标准化的评价分数,比较符合本文分析的评价方法。在对样本条约约束力的计算过程中,可以通过考察不同维度的权重占比发现样本条约存在的差异性,以此对不同条约设计的必要性进行评价分析。

熵值法的计算主要包括三个步骤。

1. 数据标准化

将各个指标的数据进行标准化处理,标准化的处理方式分为两类:

$$\text{正向指标 } X'_{ij} = \frac{X_{ij} - X_{jmin}}{X_{jmax} - X_{jmin}} \quad \text{式 1}$$

$$\text{逆向指标 } X'_{ij} = \frac{X_{jmax} - X_{ij}}{X_{jmax} - X_{jmin}} \quad \text{式 2}$$

当指标数值为越大评价越好时,选择正向指标进行数据的标准化;当指标数值为越小评价越好时,选择逆向指标进行数据的标准化。本文针对约束力的赋值设计为数值越大约束力越高,因此选用正向指标对数据进行标准化操作。

2. 求各指标信息熵

根据信息论中信息熵的定义,求得约束力评价中第 j 个内容评价的熵值 E_j ,其中:

$$E_j = -\ln(n)^{-1} \sum_{i=1}^n P_{ij} \ln P_{ij}, P_{ij} = \frac{X'_{ij}}{\sum_{i=1}^n X'_{ij}} \quad (\text{当 } P_{ij} = 0, \lim_{P_{ij} \rightarrow 0} P_{ij} \ln P_{ij} = 0) \quad \text{式 3}$$

在式3中, P_{ij} 表示针对第 j 个约束评价、第 i 个样本对指标 j 的贡献度, n 表示样本数量,常数 $\ln(n)^{-1}$ 能保证 E_j 的取值在0—1。当各样本对于评价指标 j 的贡献度完全相等时, $E_j=1$,就可以不考虑该评价指标在总体评价中的作用。

3. 计算指标权重

通过计算,可最终得到各个评价指标的权重 W_j ,其中:

$$W_j = \frac{1 - E_j}{j - \sum E_j} \quad \text{式 4}$$

当 $E_j=1$ 时, $W_j=0$,在评价过程中将不会被纳入考虑。

(五) 指标体系的应用说明

按照现实主义理论对国际关系的看法,权力是理解国际关系的关键。以关注国家安全与国际和平为宗旨的高级政治不同于以促进经济社会福利为宗旨的低级政治:高级政治所关注的是与国家权力和国家间政治相关的外交、军事和安全议题,低级政治则是指国际社会中的经济、社会和文化议题。在国际条约中,某些与高级政治相关的条约虽未涉及过多的实施细则,且通常以更加灵活松散的方式展现,但其在国际法律体系中仍然具有较高的地位和约束力度。除条约本身涉及的领域比较敏感外,更多的是因为如果国家违反高级政治领域的条约,会对自身带来更大程度的声誉损毁和国家安全威胁。^①因此,以文本约束条件得出的条约约束力值对于研究低级政治领域的条约更具现实意义。

三 约束力评价计算

在对条约约束力评价体系及计算方法进行阐述后,本部分将以中国与东盟国家签署的海洋合作条约为例,对样本内的约束力分值及特征进行分析,并据此对框架内的

^① Barbara Koremenos, "If Only Half of International Agreements Have Dispute Resolution Provisions, Which Half Needs Explaining?" *The Journal of Legal Studies*, Vol.36, No.1, 2007, pp.189-212.

约束假设进行反向验证,以确保框架的合理性和可应用性。

(一) 数据来源及筛选

本部分将分别从条约的约束力评价体系的数据来源、文本筛选、主题选择及对条约类型的选择进行说明,并对最终筛选出来的样本进行统计分析。

本文的数据来源主要包括两部分:一是联合国条约汇编。《联合国宪章》第102条规定:“一、本宪章发生效力后,联合国任何会员国所缔结之一切条约及国际协定应尽可能在秘书处登记,并由秘书处公布之。二、当事国对于未经依本条第一项规定登记之条约或国际协定,不得向联合国任何机关援引之。”^①鉴于联合国的会员国几乎涵盖所有国家及其在国际社会中的地位,它的国际条约清单在所能获取材料中是最全面的。自1946年登记公开以来,目前能在联合国条约汇编网站查询到的条约文件已达6.58万份以上,在每一份条约的记录页面中除包含条约的具体登记信息外,还包含各成员国对该条约进行签署、批准和退出等操作行为的具体日期。^②二是中华人民共和国条约数据库。该数据库由中国外交部建立,较为全面地收录了中国签署过的双边条约,是中国权威和比较完整的双边条约数据集,弥补了联合国条约汇编中未被登记的双边条约的空白,使所需数据样本尽可能全面。^③

对于条约文本的筛选,本文遵循三个原则:首先,排除一国与国际组织之间签署的条约,但会选取国际组织内谈判涉及两个或两个以上国家的条约。这是因为本文研究的主要内容是条约对国家间合作的约束力,签署一份条约至少涉及两个国家(或地区)。其次,排除无公开条约文本的条约。尽管有些条约的签署在文章或新闻报道中有所提及,但出于某种原因未能公布条约文本。本文需要根据条约细则的内容判断一份条约的约束力高低,因此没有条约文本的条约无法纳入文本分析。最后,排除未通过和未生效的条约。条约成立是条约有效性的基础,未通过及未生效的条约不构成法律责任,不在本文的研究范畴内。

在针对基本数据完成文本筛选后,本文选择中国与东盟国家签署的海洋合作条约作为指标框架的主要应用研究对象。之所以选择这一主题,是因为随着世界范围内自然资源的逐渐收紧,许多国家将国家发展合作的目标转向了海洋领域。中国在建设海洋强国的政策基础上构建了海洋政策法规体系,并不断完善和增强对海洋资源的开发能力。东盟国家多为沿海或岛屿国家,是中国的重要经贸合作伙伴和“一带一路”建设的参与者。

① 《联合国宪章·第十六章 杂项条款》, <https://www.un.org/zh/about-us/un-charter/chapter-16>, 访问时间:2021年10月18日。

② United Nation Treaty Collection, <https://treaties.un.org/Pages/AdvanceSearch.aspx?tab=UNTS&clang=en>, 访问时间:2021年10月22日。

③ 《中华人民共和国—条约数据库》, <http://treaty.mfa.gov.cn/>, 访问时间:2021年10月24日。

探究中国与东盟国家签署的海洋合作条约约束力的变化,对双方未来的海洋合作发展具有一定现实意义。此外,此类主题的条约多为合作类条约,专业性比较强、政治色彩较弱,在讨论条约约束力问题时可以最大化排除政治因素的作用。因此,本文所选取的样本仅包含海洋合作条约,而将“船舶证书”“海洋划界”等政治因素较重的非合作性条约排除在外。

一般来说,国际合作可以分为全球性合作、区域性/多边合作和双边合作,国际条约按照签署国的数量分为多边条约和双边条约。在多边条约中,有一部分条约是国际公约,是针对基本领域所确认的基本国际法,通常篇幅较长、内容较多、签署成员几乎涵盖所有国家和地区,此类条约并非因多边合作而产生的合作性条约,而是针对国际行为准则而产生的规范性条约。因此,样本中将剔除《1966年国际载重线公约》《养护大西洋金枪鱼国际公约》等海洋领域的公约。同时,因双边合作的针对性更强,条约规则相较于多边合作条约更加细致,因此在进行约束力分析时将分别对双边和多边条约进行分析,而非进行合并分析。

遵循上述原则,本文将中国与东盟国家签署的双边和多边条约信息进行提取,并在此基础上根据中国国家标准《海洋及相关产业分类》(GB/T 20794-2006)^①确定关键词,对条约内容中包含关键词的条约进行二次提取,排除其中的全球性公约,最终确定来自联合国条约汇编的8个样本(5个多边条约和3个双边条约)以及来自中华人民共和国条约数据库的22个样本(22个双边条约),共计30个有效样本。

表2显示了联合国条约汇编数据库中国所签署的条约统计信息。不难看出,中国与东盟国家的海洋多边合作相对密切,且合作内容集中在能源开采和渔业合作方面。值得指出的是,中国与东盟国家签署的海洋合作双边条约占比(2.6%)远低于中国与东盟国家签署的全部双边条约占比(8.1%),而在多边条约上的情况则相反,说明中国与东盟国家的海洋多边合作相较于其他领域的多边合作更为紧密,而与东盟各国的海洋双边合作仍有较大提升空间。此外,在可获取的中国与东盟国家签署的海洋合作双边条约文本中,部分条约的签署是基于多边条约进行的细则个性化调整。

表2 联合国条约汇编数据库记录的中国签署条约(截至2021年10月)

条约主题	条约类型	与世界各国签署的条约数量(份)	与东盟十国签署的条约数量(份)	东盟十国占比(%)
所有	双边	5110	416	8.1
所有	多边	69	46	66.7

① 新的中国国家标准《海洋及相关产业分类》(GB/T 20794-2021)已于2022年7月1日出台。

续表 2

条约主题	条约类型	与世界各国签署的条约数量(份)	与东盟十国签署的条约数量(份)	东盟十国占比(%)
海洋合作	双边	116	3	2.6
海洋合作	多边	7	5	71.4

资料来源:笔者根据联合国条约汇编整理编制。参见 United Nation Treaty Collection, https://treaties.un.org/Pages/AdvanceSearch.aspx?tab=UNTS&clang=_en, 访问时间:2021年10月22日。

注:多边条约均为多边合作/区域性条约。

从签署时间来看,中国与东盟国家海洋合作多边条约大多签署于2010年前,近10年并未有较大变动,最新的一份多边条约是在2012年生效的《南印度洋渔业协定》。中国与东盟国家签署的海洋合作双边条约在2001—2005年有较大幅度增长,此阶段共签署7份条约。另外,双边条约的合作领域在2000年前多集中于海洋运输和油气资源合作,2000年后则更加多样化,涵盖渔业、海洋运输、能源、港口建设、造船业和水资源领域。

表 3 中国与东盟国家签署的海洋合作条约情况(截至2021年10月)

国别	与中国签署的多边条约数目(份)	与中国签署的双边条约数目(份)	双边条约涉及的海洋合作领域
文莱	1	1	港口建设、能源
柬埔寨	3	0	—
印尼	2	4	港口建设、造船业、运输、渔业、能源
老挝	1	0	—
马来西亚	2	2	运输、水资源
缅甸	2	2	能源
菲律宾	4	2	能源
新加坡	1	3	运输
泰国	5	2	运输、能源
越南	3	9	运输、渔业、能源

资料来源:笔者根据联合国条约汇编和中华人民共和国条约数据库整理编制。参见 United Nation Treaty Collection, https://treaties.un.org/Pages/AdvanceSearch.aspx?tab=UNTS&clang=_en, 访问时间:2021年10月22日;《中华人民共和国—条约数据库》, <http://treaty.mfa.gov.cn/>, 访问时间:2021年10月24日。

(二) 评价结果与分析

本文对两位评价员进行了培训,使其熟悉本文建立的条约约束力框架的评价标准,并利用全变量观测法请评价员对30份样本进行打分。在本次评分中,两位评价员

在 13 个问题存在分歧,仅占比 2.2%,符合文本分析对于准确性的要求。

为探究中国与东盟国家签署的海洋合作条约中多边条约和双边条约的差异性,笔者分别对样本内的全部条约、双边条约以及多边条约进行了权重计算,结果如表 4 所示。表 4 的 1—5 列显示五个维度约束力维度的权重,6—10 列显示五个维度之下 20 个具体评价指标的权重,权重值为子指标加总值。

表 4 约束力指标权重

维度	权重 1	权重 2	权重 3	原始权重	评价指标	权重 1	权重 2	权重 3	原始权重
成员资格	0.01 ↓	0 ↓	0.03 ↓	0.11	条约类型	0.01	0	0	0.06
					成员准入条件	0.00	0	0.03	0.06
议题范围	0.02 ↓	0.02 ↓	0.09 ↑	0.06	主题细致程度	0.02	0.02	0.09	0.06
集中度	0.75 ↑	0.71 ↑	0.54 ↑	0.50	信息搜集	0.05	0.06	0.05	0.06
					信息审核	0.15	0	0.09	0.03
					审查条款	0.03	0.03	0	0.03
					术语条款	0.06	0.07	0.02	0.06
					行为规范	0.01	0.01	0.09	0.06
					行为监督	0.10	0.12	0.09	0.03
					最高惩罚方式	0.19	0.18	0	0.08
					国内法规调整	0.04	0.04	0.09	0.03
					争端解决方式	0.11	0.18	0.06	0.08
争端解决途径	0.02	0.02	0.03	0.06					
控制度	0.01 ↓	0.00 ↓	0.22 ↑	0.14	票额分配	0.01	0.00	0.05	0.06
					最低决议方式	0.01	0	0.17	0.08
灵活度	0.21 ↑	0.27 ↑	0.12 ↓	0.19	免责条款	0	0	0	0.03
					退出条件	0.03	0.03	0.02	0.06
					退出通知期	0.02	0.02	0	0.03
					退出等待期	0.03	0.04	0	0.03
					保留条款	0.13	0.18	0.09	0.06

资料来源:笔者自制。

注:权重 1 以中国与东盟国家签署的全部海洋合作条约作为计算样本,共计 30 个样本。权重 2 以中国与东盟各国签署的双边合作条约作为计算样本,共计 25 个样本。权重 3 以中国与东盟各国签署的多边合作条约作为计算样本,共计 5 个样本。原始权重为指标框架中各项指标的满分分值(即最高约束力得分)占框架总分值的比重。0 表示数值为 0,0.00 表示数值低于 0.005 但非 0。表格中箭头表示同原始权重对比下的权重上升/下降变化,表格中加粗部分表示权重较高的数值。

计算发现,中国与东盟国家签署的海洋合作双边和多边条约在成员资格、集中度方面均表现出一致的权重变化:在权重 2 和权重 3 中成员资格被给予较低权重,集中

度给予较高权重。而议题范围、控制度和灵活度的权重变化截然相反:在根据双边条约样本计算得出的权重 2 结果中,灵活度被给予较高权重,从 0.19 升至 0.27,议题范围、控制度被给予较低权重。而在根据多边条约样本计算得出的权重 3 结果中,议题范围和控制度被给予较高权重,其中控制度的增幅最大,从 0.14 升至 0.22;灵活度被给予较低权重,从 0.19 降至 0.12。同时,当样本为全部条约时,五个维度的权重变化方向与双边条约保持一致,因为双边条约样本数量多、占比大,在整体样本权重分配中起到了主导作用。

在针对双边和多边条约的每一个维度中具体指标(权重 2 和权重 3)的变化对比中发现:

在成员资格方面,条约类型和成员准入条件被给予较低权重。条约类型在全部条约中的权重(权重 1)为 0.01 而非 0,是因为此项评价在双边条约中的赋值均为 2,在多边条约中的赋值均为 0。尽管在各自样本内无差异,但将二者合并考察时则会出现差异。因此条约类型在全部条约中的权重非 0。

在议题范围方面,主题的细致程度在多边条约中被给予较高权重,是因为样本内的多边条约主题类型较为分散,而在中国与东盟国家签署的海洋合作双边条约中,主题类型一般比较单一,得分均为 1,较小的差异性给予其较低权重,因此议题范围在双边条约的权重分配中被给予比多边条约更小的权重。

在集中度方面,行为监督和国内法规调整在不同样本下均被给予较高权重,而争端解决途径也被给予较低权重。这是因为尽管在理论上争端解决存在四种强度不同的途径,但在实际形成的条约样本中,双边条约全部样本均采用内部的争端解决途径,多边条约也大多采取这一途径。同时,在权重 2 中,信息审核的权重为 0,因为在全部双边样本中,没有一项样本内条约对信息审核做出了规定;在权重 3 中,审查条款和最高惩罚方式的权重均为 0。不同的是,审查条款权重为 0 是因为全部多边条约都有审查条款,得分均为 1,而全部多边条约都未对惩罚方式做出说明,得分均为 0。此外,当权重 2 为 0.18、权重 3 为 0.16 时,则说明样本内只有一个样本有得分,其余得分均为 0。例如在争端解决方式这一评价指标中,样本内的双边条约仅有一个样本得分为 2,^①其余得分均为 0,权重 2 在争端解决方式这一指标的权重就为 0.18。总之,在集中度这一评价维度中,术语条款、行为监督、最高惩罚方式和争端解决方式在权重 2 中被给予较高权重,而信息审核、行为规范、行为监督和国内法规调整在权重 3 中也被给予较高权重。

^① 1962 年《中华人民共和国和越南民主共和国通商航海条约》中明确规定“若发生争执……由专门设定的仲裁法庭审理该项争执”,因此得分为 2,其余双边条约均为友好协商解决争端,得分全部为 0。

在控制度方面,权重2中的票额分配和最低决议方式指标的权重均接近或等于0。根据样本评价规则,双边条约中几乎所有条约都会包含“平等互利原则”字样,^①且因为达成协议的主体只有2个,必须全票通过才可形成条约,因此双边条约在这一指标中通常会获得指标的满分。多边条约因为签署国众多,一般都会对票额分配和最低决议方式做出说明,但在样本内的5个多边条约中,针对票额分配有3份条约明确表示“每个成员国有一票”,有2份条约未做出说明;而最低决议方式除《亚洲打击海盗和持械抢劫船只区域合作协定》中“应以协商一致方式做出决定”的规则获得最高分值3以外,其他多边条约的最低决议方式均为“绝对多数/简单多数”。

在灵活度方面,样本内双边和多边条约均没有对免责条款做出说明,得分均为1,因此无论是总体、双边还是多边样本内,给予这一指标的权重均为0。多边条约一般会对退出通知期和退出等待期做出明确说明,且约束力评价得分相同,仅在保留条款和退出条件上略有不同。在双边条约中,部分样本虽然没有明确规定条约的退出等待期,但也在条款中说明其“终止不影响已经达成但未完成的项目”,本文也将这种情况视为存在退出等待期,因此双边条约样本在对灵活度的条约约束评价上存在一定差异。

通过针对不同样本权重计算结果的分析,本文认为双边和多边条约在文本设计上确实存在差异,简单合并计算并不能得到较为客观的权重分配和约束力评价得分。因此,本文将双边和多边条约进行分开计算和分析:在计算双边条约约束力评价得分时采用权重2,在计算多边条约约束力评价得分时采用权重3。重新分配权重可以得出中国与东盟国家签署的海洋合作条约的约束力评价最终得分,此时双边和多边条约所保留的评价指标出现了不一致。下文对多边条约和双边条约的指标筛选结果、约束力评价得分特征值、各维度约束评价得分表现和约束评价总特征进行逐一阐述。

多边条约最终保留了五个维度的14个评价指标,删除了条约类型、审查条款、最高惩罚方式、免责条款、退出通知期和退出等待期指标。在表5中,五个多边条约约束力评价总得分的平均值为0.42、最大值为0.59、最小值为0.14,标准差为0.18。其中,约束力评价得分最高的条约是《建立印度洋金枪鱼委员会协定》。在五个维度的约束力评价得分分布中,控制度的权重占到总分的22%,但是其均值相对较低,说明在样本内的大部分多边条约中控制度的约束较低;在集中度的8个指标中,超过半数的多边条约对信息审核、行为监督和国内法规调整没有提出要求,且在行为规范的约束方面也较为模糊;灵活度约束评价表现最佳,说明相较于其他约束条件,多边条约对于灵

^① 仅《中华人民共和国政府和泰国政府关于胜利原油长期贸易议定书》和2005年《中华人民共和国政府和新加坡共和国政府海运协定》两份条约文本中未出现“平等”字样。

活条款的设计更具约束性;这些条约几乎均对退出条件做出了“在规定时间范围”的期限约束。此外,在每一项中国与东盟国家签署的海洋合作多边条约中,都至少存在一个维度的约束力评价得分为0。据此推测,在多边条约的文本设计中,可能存在五个维度在约束设计方面相互替代的关系。

表5 中国与东盟国家签署的海洋合作多边条约约束力评价得分

条约名称/评价指标	总体	成员资格	议题范围	集中度	控制度	灵活度
《建立印度洋金枪鱼委员会协定》	0.59	0.01	0	0.40	0.05	0.12
《亚洲打击海盗和持械抢劫船只区域合作协定》	0.54	0	0.09	0.28	0.17	0
《南印度洋渔业协定》	0.54	0.03	0.09	0.30	0	0.12
《关于亚洲和太平洋水产养殖中心网络的协定》	0.28	0.01	0	0.19	0.05	0.02
《关于设立印度—太平洋渔业理事会的协定》	0.14	0.01	0	0.05	0.05	0.02
分配权重	1	0.03	0.09	0.54	0.22	0.12
平均值	0.42	0.01	0.04	0.24	0.06	0.06
标准差	0.18	0.01	0.05	0.12	0.05	0.05

资料来源:笔者自制。

双边条约最终保留了三个维度的15个评价指标,删除了条约类型、成员准入条件、信息审核、最低决议方式和免责条款指标。在表6中,样本内双边条约约束力的平均得分为0.18、最大值为0.43、最小值为0.04,标准差为0.11。在三个维度的约束力评价得分中,议题范围的平均值较多边条约要小很多,并非双边条约的主题更为笼统;相反,双边条约的议题范围多集中在高分区域,但因其差异性低而导致所分配的权重低,使得双边条约在此项的得分比多边条约低。而在包含了9个指标的集中度约束评价中,超过半数的双边条约未对信息搜集和术语条款做出要求;而在行为监督、最高惩罚方式和争端解决方式指标中,只有极个别的条约对此做出了约束规范,这使得几乎全部双边条约在这些约束指标的约束力评价得分均为0。其中,仅《中华人民共和国政府和越南社会主义共和国政府关于北仑河口自由航行区航行的协定》《中华人民共和国政府和越南社会主义共和国政府北部湾渔业合作协定》《印度尼西亚共和国交通部和中华人民共和国交通部交通领域技术合作谅解备忘录》对行为监督提出明确要求,规定设立专门的工作组或协会监督条约执行内容;仅《中华人民共和国政府和越南社会主义共和国政府北部湾渔业合作协定》对最高惩罚方式提出明确要求,指出“缔约各方有权根据各自国内法……进行处罚”。在灵活度中,虽然约束力评价得分最大值

为0.20,但是平均值仅为0.05,说明大部分双边条约对于退出机制的规定相较于多边条约都更为松散。在多边条约样本中,均对退出机制做出了详细规定,且退出通知期一般为12个月,也有少部分多边条约规定为6个月或3个月。而退出等待期一般以年为单位且通常在2年以上,也有少部分为半年。在双边条约中,仅有不到70%的样本对退出通知期做出了规定,且退出通知期一般以6个月居多,此外,不到50%的双边条约对退出等待期做出了规定,等待期一般以项目进度为标准。在对保留条款的规定中,双边条约中仅1962年《中华人民共和国和越南民主共和国通商航海条约》对保留提出了明确要求,指出“为了国家安全……双方得保留对……规定和限制禁止的权利”。

在去掉因双边条约的固有性质定义的指标后,^①双边条约在议题范围、集中度和灵活度方面的约束表现均不如多边条约,且双边条约的约束力评价得分的标准差比多边条约更低,说明样本内双边条约的约束力评价得分较多边条约的约束力评价得分分散程度更低、条约文本设计的差异性更低。

表6 中国与东盟国家签署的海洋合作双边条约约束力评价得分分析

约束力评价	总体	议题范围	集中度	灵活度
样本数	25	26	26	26
分配权重	1	0.02	0.71	0.27
平均值	0.18	0.01	0.12	0.05
最大值	0.43	0.02	0.42	0.20
最小值	0.04	0	0.01	0
标准差	0.11	0.01	0.09	0.04

资料来源:笔者自制。

在分别讨论了双边和多边条约的条约约束情况后,可以将国别条约约束差异分析纳入考量。表7反映了中国与东盟国家样本内签署的海洋合作多边条约的约束力评价得分情况。其中,中国与文莱、老挝、新加坡签署的多边条约约束力评价得分的均值最高(0.54),与越南、柬埔寨签署的多边条约约束力评价得分的均值最低(0.32)。按照条约签署的年份,东盟国家与中国签署的多边条约的约束力评价得分整体呈现逐年递增的态势(如图1),说明中国与东盟国家签署的海洋合作多边条约的文本设计越来越规范严谨。此外,中国与东盟国家签署的每一份多边条约都至少涉及4个东盟国家,且随着东

① 条约类型、成员准入条件、票额分配和最低决议方式指标。

盟国家越来越统一的集体行为,在签署多边条约的过程中包含的东盟成员国数量也越来越,中国与不同东盟国家签署的多边条约约束力评价得分由此逐渐趋于一致。

表 7 中国与东盟国家签署的海洋合作多边条约约束力评价得分对比

国别	条约数量	约束力累计值	约束力最大值	约束力最小值	约束力均值	约束力标准差
文莱	1	0.54	0.54	0.54	0.54	0
柬埔寨	3	0.96	0.54	0.14	0.32	0.16
印尼	2	0.73	0.59	0.14	0.37	0.22
老挝	1	0.54	0.54	0.54	0.54	0
马来西亚	2	0.87	0.59	0.28	0.43	0.15
缅甸	2	0.82	0.54	0.28	0.41	0.13
菲律宾	4	1.55	0.59	0.14	0.39	0.18
新加坡	1	0.54	0.54	0.54	0.54	0
泰国	5	2.09	0.59	0.14	0.42	0.18
越南	3	0.96	0.54	0.14	0.32	0.16

资料来源:笔者自制。

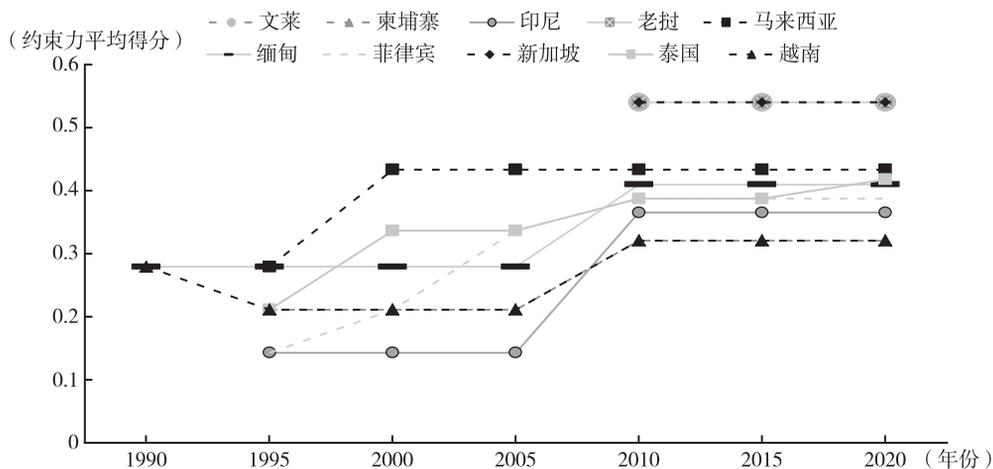


图 1 1990—2020 年中国与东盟国家签署的生效海洋合作多边条约约束力平均得分变化

资料来源:笔者自制。

表 8 反映了中国与东盟国家签署的海洋合作双边条约的约束力评价得分情况。其中,中国与越南签署的双边条约约束力评价得分的均值最高(0.23),与菲律宾签署

的双边条约约束力评价得分的均值最低(0.04)。从数量上看,中国与越南签署的双边条约最多(九份),条约约束力评价得分最大值为0.43、最小值为0.09,标准差为0.13。这说明与其他国家相比,中国与越南签署的双边条约的差异更大,条约约束力评价得分多集中在极值两端,但总体来看这些条约仍处于较高约束水平。中国与印尼签署了四份双边条约,涉猎主题较为全面,而条约约束力评价得分最大值为0.27、最小值为0.05,标准差为0.08。可见,虽然中国与印尼签署的双边条约所涉及的海洋合作领域较为广泛,但约束力评价得分都不高。在本文的指标框架中,中国与柬埔寨、老挝签署的双边条约均为0份。其中柬埔寨是因为样本没有海洋合作的可公开条约文本,而老挝作为东盟唯一的内陆国与中国自然不会有与海洋相关的双边合作。按照条约签署的年份,除越南、文莱和菲律宾外,中国与东盟国家签署的双边条约约束力均呈逐步上升趋势(如图2)。菲律宾之所以在1990—2015年的约束力评价得分降至0,是因为样本内中菲于1979年签署的唯一一份海洋合作双边条约《中华人民共和国政府和菲律宾共和国政府长期贸易协定》在1985年失效后,直到2018年两国才又签署了另一份合作条约《中华人民共和国政府和菲律宾共和国政府关于油气开发合作的谅解备忘录》。在中国与越南签署的双边条约中,虽然在2005年的统计中出现了一个极端值0.30,但此后的条约评价得分均较低并呈现逐年下降的趋势。而在2005年的统计节点出现极端值,是因为在2004年中国与越南签署了《中华人民共和国政府和越南社会主义共和国政府北部湾渔业合作协定》,这份条约在集中度和灵活度的约束力评价得分中都属于较高的水平,约束力评价得分总值为0.43,因此拉高了均值。

表8 中国与东盟国家签署的海洋合作双边条约约束力评价得分对比

国别	条约数量(份)	约束力累计值	约束力最大值	约束力最小值	约束力均值	约束力标准差
文莱	1	0.14	0.14	0.14	0.14	0.00
柬埔寨	0	0	0	0	0	—
印尼	4	0.71	0.27	0.05	0.18	0.08
老挝	0	0	0	0	0	—
马来西亚	2	0.37	0.20	0.17	0.18	0.02
缅甸	2	0.19	0.11	0.08	0.09	0.02
菲律宾	2	0.09	0.05	0.04	0.04	0.00
新加坡	3	0.57	0.24	0.13	0.19	0.04
泰国	2	0.34	0.24	0.11	0.17	0.06
越南	9	2.10	0.43	0.09	0.23	0.13

资料来源:笔者自制。

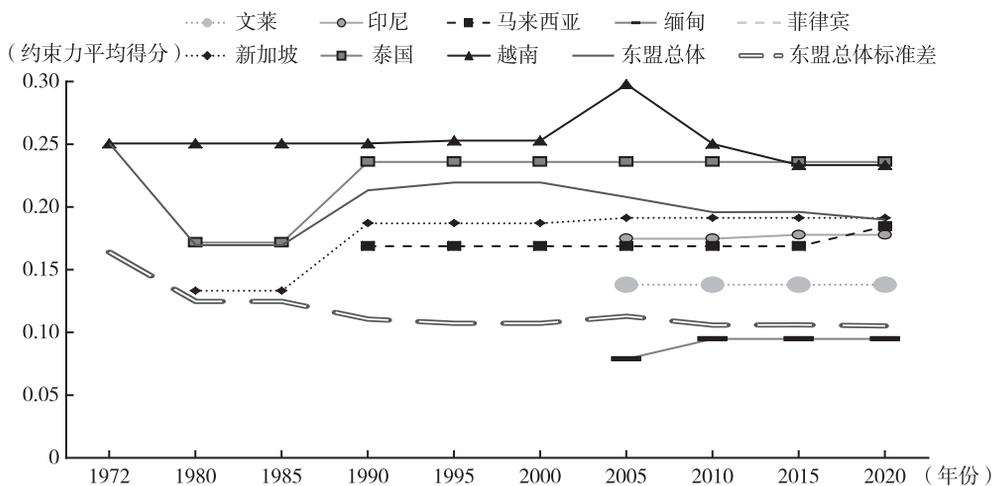


图2 1972—2020年中国与东盟国家签署的生效海洋合作双边条约约束力评价得分变化
资料来源:笔者自制。

总体来看,中国与东盟国家签署的海洋合作双边条约中,约束力评价得分均值在2000年以后均呈现逐年下降趋势,且约束力评价得分的标准差的年度变化从1972年的0.12降至2020年的0.10,这说明中国与东盟国家签署的海洋合作双边条约的约束差异在逐渐缩小。

四 结论与讨论

基于凯里迈诺斯对国际条款设计规则与国家间合作关系的研究成果,本文构建了一套相对系统且具体的条约约束力评价体系。这一框架体系分为五个维度、20个指标。笔者在指标框架基础上通过文本分析的全指标分析法和熵值法对条约约束力进行计算,以中国与东盟国家签署的海洋合作条约作为研究对象,分别从指标权重分配、约束力评价得分、国别对比以及得分均值年度变化四方面进行了一系列实证分析,发现中国与东盟国家签署的海洋合作多边条约与海洋合作双边条约在条约文本设计上存在较大差异,通过熵值法计算得出的双边和多边条约在各评价指标的不同权重分配结果也验证了这一观点。此外,中国与东盟国家签署的双边条约和多边条约在条约文本设计和年度约束力评价得分变化上也有不同表现。

同时,在实证检验中也发现框架内的某些指标设定仍存在问题。在权重分配计算结果中针对多边条约样本的条约类型、审查条款、最高惩罚方式、免责条款、退出通知

期和退出等待期六项评价指标被完全剔除;在针对双边条约样本权重分配计算结果中,条约类型、成员准入条件、信息审核、最低决议方式和免责条款5项评价指标被完全剔除。这些被剔除的指标集中在成员资格、控制度和灵活度三个维度上:一是成员资格方面。首先,在双边和多边条约的权重分配中均剔除了条约类型。原因在于,在约束框架中条约类型是基于联合国条约汇编对于条约文本三种分类进行的定义,将条约文本分为双边条约、开放多边条约和封闭多边条约,且样本内的多边条约均属于开放多边条约。这就使得在最初的约束力赋值中,双边条约得分均为2,多边条约得分均为0,导致根据熵值法计算的样本内权重分配值均为0。条约类型设计的初衷是希望在理论上赋予双边条约更强的约束值,因为国家数量的多寡确实对条约约束有着切实的影响,但在实际操作中这一设定存在缺陷。当默认双边和多边条约在设计中存在差异时,不能将二者一概而论,否则这一指标将失去意义。其次,在针对样本筛选的过程中,本文将国际性条约和区域性条约进行了区分,在未来的框架完善中也可考虑根据地缘特征和经济特征等将成员准入条件进行细分。最后,在最终的约束力评价得分计算中,双边条约的约束力评价得分比多边条约的约束力评价得分要更低,显然不符合之前的理论假设,这说明在成员资格这一维度上的评价依据还有待完善。二是控制度方面。审查条款可能在设计上存在一定不足,是因为在样本内全部多边条约均存在有关审查大会的条款,未来可以考虑细化审查大会的指标评价层级,如规定每年、每两年、每五年、不定期等不同的召开时间和条件,进行约束力差异的研究和层级划分。剔除最高惩罚方式则主要是因为针对特定样本的研究中均没有对最高惩罚方式的说明,这并不能否定惩罚条款对条约约束所带来的作用(免责条款、信息审核也因此原因被去掉)。三是灵活度方面。在多边条约样本内的权重分配中,退出通知期和退出等待期被剔除是因为其约束力评价得分完全一致。在本文现有研究框架下,二者的约束等级差异均只分为“有”和“无”,而在多边条约的条款设计中,通常都会包含对于此两项指标更加具体的说明,未来在研究中可基于不同具体期限条件对退出通知期和退出等待期做进一步区分。

本文旨在展现条约约束力评价体系的可应用性,在针对所选主题的样本数量上可能存在一定不足,如多边条约出现数量偏少和条约可能存在遗漏等问题。为保证分析的全面性,本文通过统一数据源、统一数据筛选规则尽可能使可用条约样本最大化。尽管本文建立的约束力评价体系有待进一步完善,但随着学者对评价体系的探索逐步成熟,无论是对于不同主题、不同国家之间的国际条约设计特征研究,还是对于国家间签署同类型条约的不同约束选择研究都具有一定现实意义。