

**A FRAMEWORK FOR EVALUATING
THE PERFORMANCE OF INTERNATIONAL COURTS AND TRIBUNALS**

THERESA SQUATRITO, ORAN R. YOUNG, ANDREAS FOLLESDAL, AND GEIR ULFSTEIN
APPROXIMATELY AS APPEARS IN
The Performance of International Courts and Tribunals,

THERESA SQUATRITO, ORAN R. YOUNG, ANDREAS FOLLESDAL and GEIR ULFSTEIN (EDS),
CAMBRIDGE UNIVERSITY PRESS 2018, 3-35

Introduction

Many analysts have noted a trend toward the legalization of international affairs in recent decades (Goldstein et al. 2000). While the extent of this development is hard to measure, one significant feature of legalization is unmistakable. We have witnessed a proliferation of international courts and tribunals (ICs).¹ Such bodies now operate globally and in several regions of the world; they play significant roles in the application and interpretation of many elements of international law. ICs address issues of regional integration, trade, and economic relations. They have become important elements in human rights systems in Europe, the Americas, and Africa. ICs are integral to the development of international criminal law and help to determine territorial and maritime boundaries.

There is significant variation among ICs regarding not only their mandates, but also the practices they have adopted and the effects of their rulings. International courts vary greatly with regard to their level of activity. Some courts are extremely busy institutions. The European Court of Human Rights (ECtHR) and the Court of Justice of the European Union (CJEU)² both issue hundreds of decisions annually. The ECtHR issued 916 judgments responding to 3,659

¹ We use the abbreviation ICs rather than ICTs to avoid confusion with international criminal tribunals. But we cast a broad net in defining our universe of cases.

² Formerly known as the European Court of Justice (ECJ).

applications in 2013, while the CJEU issued 1322 decisions by judgment or order in 2013.³ On the other hand, there are courts like the International Court of Justice (ICJ) that hear only a handful of cases in any given year. ICs differ also in their practices pertaining to access, transparency and confidentiality, fact-finding, standards of review, methods of interpretation, and more.

Some ICs have become “islands” of litigation, generally addressing a narrower set of legal questions than their mandates cover. Approximately ninety percent of the cases on the docket of the Andean Tribunal of Justice, for example, pertain to intellectual property law, though the mandate of the court is much broader (Helfer et al. 2009). On the other hand, there are courts that have exercised authority across the entire scope of their mandate; some have even expanded their jurisdiction. Similarly, some ICs have demonstrated judicial activism, while others have shown judicial restraint.

Several international courts have encountered serious problems, facing backlogs in their caseloads, backlash on the part of member states, or failure to become operational. The ECtHR, for example, is saddled with an overwhelming backlog of cases; it had a backlog of 99,900 applications in 2013. The Southern African Development Community (SADC) Tribunal’s operations were suspended as a result of backlash on the part of Zimbabwe. Twenty ICs have never become operational or became defunct after deciding only a few cases (Romano 2014).

ICs have made varying contributions to global governance. First, ICs often contribute to state compliance with international legal commitments, though at times they have sparked defiance, as in the case of the SADC Tribunal. Some ICs have been able to influence the behavior of states, encouraging them to adjust domestic laws to comply with international law. Others seek remedies by identifying appropriate compensatory measures for non-compliance.

³ The figures for the CJEU include the decisions of the General Court and the Court of Justice. For statistics on the ECtHR’s judicial activity, see www.echr.coe.int/Pages/home.aspx?p=reports&c= (accessed 27 July 2017). For statistics on the judicial activity of the CJEU, see http://curia.europa.eu/jcms/jcms/Jo2_7000/ (accessed 27 July 2017).

Second, some ICs have influenced the establishment and the work of other ICs, while others have not had such effects. The International Criminal Court (ICC), for instance, has been shaped by the International Criminal Tribunal for the former Yugoslavia (ICTY) and the International Criminal Tribunal for Rwanda (ICTR). Third, ICs have contributed to broader processes of global governance. The CJEU, for example, has been central to the institutionalization of regional integration in Europe (Burley and Mattli 1993; Mattli and Slaughter 1998; Stone Sweet 2004); the Dispute Settlement Mechanism (DSM) of the World Trade Organization (WTO) has facilitated trade relations globally (Rosendorff 2005; Goldstein et al. 2007; Davis 2012). Evidence has established a link between international criminal prosecutions and processes of democratization (Sikkink 2011). Some attribute the judicialization of international affairs and constitutionalization of international law to the experiences of the CJEU, ECtHR, the WTO DSM, and others.

As a matter of fact, then, there is a great deal of variation in the *performance* of international courts. This observation raises several critical questions. How should we think about performance in this context? How can we measure performance in specific cases? Why do some ICs perform better than others? What are the determinants of the performance of these courts? Are there ways to improve the performance of international courts? These are the concerns that motivate the contributions to this book.

The study of judicial performance

While prior studies have tended to focus on specific questions relating either to the design or to the effects of international courts, we develop an integrated framework for the study of the performance of ICs. Using this framework, the contributors to this volume present empirical assessments of the performance of international courts that consider both the results that courts

produce and the procedures guiding their operation. We explore factors that may explain the patterns of performance we observe.

This study of IC performance takes a broad comparative approach covering the full array of international courts and tribunals. We define the universe of cases to encompass international judicial bodies that: (1) decide the question(s) brought before them on the basis of international law, (2) follow pre-determined rules of procedure, (3) issue legally binding decisions or opinions, (4) are composed of independent members, and (5) require that at least one party to a dispute be a state or an international organization (Romano et al. 2014: 6). While some analysts may argue that international criminal tribunals do not fulfill the fifth criterion, the office of a Prosecutor, an organ of an international organization, is one of the parties to disputes in international criminal tribunals (Romano et al. 2014: 7). While ICs can issue non-binding opinions (i.e., advisory opinions) in addition to legally binding decisions, this definition excludes quasi-judicial bodies, such as the United Nations human rights treaty bodies, which issue only non-binding recommendations. In addition to permanent bodies, we include ad hoc judicial bodies that meet these criteria, such as investor-state arbitration tribunals.⁴ This allows us to assess the extent to which permanence contributes to performance. We include a range of ICs spanning several issue areas, including human rights, trade, investment, and criminal law.

We are interested both in the outcomes courts produce and the processes through which they arrive at judgments. Drawing on research dealing with the performance of international organizations (Gutner and Thompson 2010), we label these dimensions *outcome performance* and *process performance*. Outcome performance refers to the degree to which ICs attain substantive goals. Process performance, on the other hand, is a matter of the degree to which IC practices measure up to intended or aspired procedural standards.

⁴ Permanence does not refer to whether the court or tribunal itself is permanent. Instead, it means that “they are made of a group of judges who are sitting permanently and are not selected ad hoc by the parties for any given case” (Romano 2011: 262).

There are several reasons to study the performance of international courts. For one thing, the performance of these courts affects their legitimacy. As Buchanan and Keohane argue, “If an institution exhibits a pattern of egregious disparity between its actual performance, on the one hand, and its self-proclaimed procedures or major goals, on the other, its legitimacy is seriously called into question” (2006: 422). A court's performance may also affect levels of popular support or the politicization of international legal processes (for example, see Helfer and Alter 2013). Understanding the performance of ICs has implications as well for the design and reform of international courts. Knowing the determinants of good and bad performance can help in identifying what reform efforts are merited and how reform should proceed. Observers of the European Court of Human Rights, for example, have noted that reform efforts have been informed by policy-relevant assessments of the Court's performance. In addition, the performance of ICs can have implications for the effectiveness of international regimes or governance systems in which they are embedded. Thus, studying IC performance may advance our understanding of effectiveness of international regimes or governance systems.

The study of performance draws on existing research on the performance of international organizations (Gutner and Thompson 2010; Tallberg et al. 2016) and international environmental institutions (Mitchell 2008) as well as research on the effectiveness of international courts (Shany 2014) and international regimes or governance systems (Young 1999; Miles et al. 2002; Hovi et al. 2003; Breitmeier et al. 2006). The literature on international regimes, governance systems, and social institutions more broadly has focused on the extent to which these arrangements contribute to solving problems (e.g., the depletion of stratospheric ozone, the occurrence of genocide) or steering systems toward socially desirable outcomes (e.g., increased trade, enhanced respect for human rights). The term effectiveness is used in this literature as a measure of the success of regimes and governance systems in solving problems or moving systems toward desired outcomes.

International courts, in our view, are tools or mechanisms that play a number of roles in regimes or governance systems. It is worth differentiating among several distinct roles in this regard. First, courts play roles within individual regimes or governance systems by providing authoritative interpretations of the meaning of a regime's rules, adapting rules to new circumstances, resolving disagreements among parties regarding the meaning of the rules, and helping to ensure compliance on the part of those subject to a regime's rules. Second, courts help to sort out tensions or conflicts between individual regimes or governance systems. This is typically a matter of determining the spheres of applicability of different rules or determining which rules take precedence over others in cases of direct conflict. Third, courts play a broader role in ensuring that the operations of international regimes or governance systems conform to overarching principles, norms, and values applicable above the level of individual regimes or governance systems. In many cases, this is a matter of ensuring that individual regimes or governance systems adhere to procedural norms like those associated with the idea of fairness.

To know whether regimes or governance systems are effective, then, we argue that it is imperative to ask whether international courts perform their roles well or poorly. The existence of courts that perform well can and often will contribute to the effectiveness of international regimes or governance systems. But there is nothing automatic about this relationship. It is possible to imagine regimes that are successful in solving various problems even in the absence of courts that perform well. Conversely, we can imagine regimes that fail to solve problems, despite the existence of courts that perform their roles well. The relationship between the effectiveness of international regimes or governance systems and the performance of international courts is a matter that merits careful analysis.

Assessing the performance of international courts, on this account, is a matter that extends beyond the intentions of those who create them. Creators may or may not articulate the roles they expect international courts to play. Moreover, creators may not agree on what roles

courts are intended to have. Whether or not the creators are explicit about these matters, however, we can ask questions about the performance of a court in interpreting and adapting rules, sorting out tensions among different regimes, and encouraging adherence to broader principles and norms. As Mitchell observes in his general account of institutional performance, “performance analysis seeks to identify how much an institution contributed to whatever progress was made toward a specified goal ” (2008: 79). But as he goes on to say, the relevant goals may be specified by “the creators of the institution, other interested parties, or the evaluator” rather than reflecting only the intentions of the creators (79). In this sense, our focus on performance is similar to Shany’s analysis of the “effectiveness” of international courts comparing “...actual impacts with desired outcomes, or performance with expectations ... in the eyes of multiple constituencies” (2014: 6). Unlike Shany, however, we employ the concept of performance to capture a broad set of goals identified by analysts as well as relevant constituencies.

There is an inescapable normative dimension to any study of the performance of international courts. A study of the performance of ICs that looks beyond the intentions of the creators is compatible with evidence suggesting that the effects of these courts go well beyond initial intentions. As Caron argues: “When assessing the value ... of international courts and tribunals scholars should not only proceed in terms of how well a given institution serves its constituted ends, but also how well it serves the unstated purposes” (2006: 410). More generally, there are normatively grounded differences among observers regarding the proper interpretation of a regime’s rules, the precedence granted to one regime vis-à-vis another, and the broader principles to be applied to the operations of regimes. Those who agree that a court performs well, therefore, may be more or less satisfied with the results in normative terms. But this should not detract from efforts to determine how well international courts play their roles and to identify factors that explain variance in performance.

Assessing Performance

Performance as an explanandum is a multifaceted concept. In everyday usage it has two distinct but related meanings. First, as a verb, to perform is simply to fulfill an obligation or complete a task. Second, as a noun, performance refers to the manner in which a task is completed. Thus to address the issue of performance, as applied to the social world, is to address both the outcomes produced and the process—the effort, efficiency and skill—by which goals are pursued by an individual or organization (Gutner and Thompson 2010: 231).

We build on this formulation, adopting Gutner and Thompson's distinction between outcomes and processes and adapting it for application to ICs. Thus, one dimension of performance is *outcome performance*, covering the full range of outcomes that result, either directly or indirectly, from the operation of a court.⁵ The second dimension is *process performance* associated with the way in which international courts exercise their authority.

To assess both dimensions of performance, we need criteria against which outcomes and processes are compared. As Gutner and Thompson observe, "[e]stablishing a baseline is important because it is only against a particular set of objectives and in the context of a given timeframe that performance can be assessed" (2010: 240). But the selection of appropriate criteria against which performance is compared depends not only on the mandate of a court but also on the perspective of the researcher. Thus, Mitchell notes that while "[i]nstitutions can be

⁵ We use outcome here as a generic term to encompass a broad set of results. For these purposes, we do not distinguish outcomes from outputs and impacts, as is typical in comparative politics literature or in regime effectiveness literature. In other words, the outputs of an international court (or their judgments) are among the range of outcomes resulting from the IC's operation. The same applies to impacts.

evaluated against either the primary or the subsidiary goals for which they were designed...they can also be evaluated against the goals of actors outside an institution in question” (2008: 80).

It follows that the criteria of evaluation can vary considerably. Still, we endeavor to develop realistic criteria based on common understandings of international courts. We start by identifying the functions or roles that analysts typically ascribe to international courts. Some of these are clearly intended by creators of ICs, as reflected in their constitutive documents. Others are roles that ICs serve in practice, irrespective of creators’ intentions. We then posit that these are plausible criteria against which the performance of courts can be evaluated. We do not claim that these criteria apply equally to all courts. Nor do we assume that they are exhaustive. Rather, our intention is to provide a common analytical starting point for comparison across ICs. We recognize that it is necessary to contextualize criteria on a case-by-case basis, which all the contributions to this volume do. We also recognize that courts rarely bear sole responsibility for the production of specific outcomes. Our objective is to make judgments regarding the proportion of the variance in outcomes that we can attribute to the activities of specific courts.

The analysis of *outcome performance* is a matter of assessing the results of court activities. This means evaluating the extent to which the actions of courts contribute to the attainment of substantive goals. The relevant goals may be framed at the level of specific cases, the level of the governance system to which cases belong, or the level of the development of international society as a whole. Thus, the relevant criteria may range from a narrow conception of performance to much broader conceptions. Drawing on existing literature, however, we approach outcome performance in terms of the roles or functions of courts in various social settings. Specifically, we identify criteria dealing with dispute settlement, clarification of the law, and compliance.

The first criterion involves the function of settling disputes. To what extent are courts able to settle disputes? How effectively do they do so? Extensive research suggests that

international courts and tribunals do contribute to the settlement of disputes (Alter 2013; Bogdandy and Venzke 2013; Alvarez 2014; Shany 2014). Traditional perspectives on the dispute settlement function of ICs focus on the settlement of disputes between states. More recently, several ICs have begun to address disputes between private parties as well as states. Human rights courts, for instance, often have jurisdiction over disputes between private actors and states. International organizations also may be involved in cases that come before ICs, as in cases where the UN General Assembly requests advisory opinions from the ICJ, or when the CJEU reviews disputes between the Commission of the European Union and a member state. While dispute settlement is generally recognized as an important role of ICs, it is not a criterion that applies to all courts. Many scholars conclude that international criminal courts or tribunals prosecute crimes rather than acting to settle disputes (Alter 2013; Bogdandy and Venzke 2013). Still, dispute settlement is an important focus in our examination of the performance of ICs. Are they good at settling disputes? What enables courts to settle disputes effectively?

A second criterion for assessing the outcome performance of international courts is clarification of the law. Most treaties establishing courts assign them the task of “interpreting” the law. Several scholars also emphasize the importance of clarifying the law (Alter and Helfer 2010; Bogdandy and Venzke 2013; Alvarez 2014; Helfer 2014).⁶ Law, including international law, comprises abstract rules prescribing behavior. These rules provide general prescriptions; they cannot anticipate all possible circumstances to which they will apply. Also, states may disagree on the content of rules and thus intentionally adopt imprecise language in legal agreements. The result is that legal rules are often incomplete and ambiguous with regard to their application. ICs frequently are called upon to interpret the law and to clarify how general prescriptions apply to specific circumstances.

⁶ Not all scholars use the terminology of *clarification*. They may speak of lawmaking, gap-filling, developing the law, or other similar terms.

Clarifying the law can take a number of forms. International courts may interpret the scope of legal rules. The ICJ and the ICTY, for example, have redefined the scope of laws of war to apply to internal conflicts (Haye 2008). Clarifying the law also may require ICs to order or prioritize different rules. A central question frequently brought before courts involves determining whether specific rules are superior or subordinate to one another. Similarly, ordering occurs when an international court determines whether a law is *jus cogens* or customary law. By determining that a law such as the prohibition of torture is *jus cogens*, the court is asserting that the prohibition of torture is a supreme law. In such cases, courts clarify the law by determining the hierarchy of rules. Beyond this, ICs may determine how rules balance against one another in order to reconcile conflicting rules. This may entail the balancing of rules within a single legal regime or the balancing of rules across legal regimes. An instance of the former involves the efforts of the ECtHR to clarify the law by determining how to balance the rights of one individual against those of another or against societal interests. The latter arises when courts help to adjudicate conflict between international trade law and international environmental law, as in the judgment of the WTO's Appellate Body (AB) in the *Shrimp-Turtle* case. The clarification of the law sometimes involves very limited adjustments to legal rules. At other times, clarification may lead to extensive revisions in the law, causing some observers to speak of the role of courts in developing the law, lawmaking, or even in constitutionalizing international law.

A third criterion for evaluating outcome performance centers on the expectation that ICs facilitate compliance with international law. Some scholars refer to international courts as having enforcement or compliance functions (Shelton 2009; Alter 2013) or even treat compliance as a measure of performance (Guzman 2008).⁷ ICs can play this role either by providing incentives to comply with legal commitments or by providing remedies in response

⁷ For a discussion of the latter, see Shany (2014).

to noncompliance. Courts may encourage compliance through a process of social sanctioning, in effect delegitimizing actors or behaviors that do not comply with community norms or social expectations. Social sanctioning by ICs can enhance compliance through three distinct mechanisms. Such sanctions can damage an actor's reputation, causing a loss of status. Alternatively, the judgments of an IC may encourage more powerful actors, such as states or in some instances international organizations, to exercise their power to compel compliance. These more powerful actors may use the ruling of an international court to enhance the legitimacy of their exercise of power. Such exercises of power may range from the distribution of aid and trade benefits to benefits pertaining to membership and positions within international organizations and even to retaliatory measures and military actions. A third (and arguably most significant) mechanism comes into play when IC decisions strengthen domestic compliance constituencies, including domestic electorates and NGOs. Such constituencies may hold the executive to account for violations of international law by shifting political support to the opposition party or mobilizing social groups to demand domestic policy changes. Research suggests that all these mechanisms make a difference in specific cases (Dai 2005; Simmons 2009; Hafner-Burton 2013; Helfer and Voeten 2014).

In some cases, international courts may be able to influence compliance directly, through the imposition of penalties and other forms of punishment and through the monitoring of implementation. Some ICs have the authority to impose monetary penalties. The CJEU, for example, can impose penalties on states in response to noncompliance. ICs may also be able to require that compensation be paid to a wronged party. For example, the Inter-American Court of Human Rights (IACtHR) has authority to require monetary compensation be given to victims, which may incentivize states to adapt their behavior to comply with applicable laws. Monetary penalties can also include the authorization of retaliatory measures that have economic consequences. The WTO's Dispute Settlement Understanding, for instance, allows a state to

use countermeasures against another state that fails to comply with the ruling of a panel or the Appellate Body. Alternatively, some ICs have the authority to impose sanctions in the form of physical punishment. International criminal courts can impose incarceration on guilty parties. Also, Article 94(2) of the UN Charter authorizes the Security Council to take measures (including the use of force) against states that fail to comply with ICJ decisions, thereby threatening peace and security (Bogdandy and Venzke 2013).⁸

Beyond this, an IC may enhance incentives to comply by initiating a process of monitoring the implementation of its decisions, often through an international body linked to the court. When the ECtHR finds that a law has been violated, for instance, the Committee of Ministers of the Council of Europe monitors the implementation of the Court's judgment. The WTO's Dispute Settlement Body plays a similar role in monitoring compliance with decisions of panels and the AB. In a more decentralized process, monitoring of implementation also may occur through actions of domestic entities. As in the case of social sanctioning, it is unclear exactly how effective these mechanisms are at influencing compliance. Further research regarding all these mechanisms should help to illuminate the performance of ICs.

Process performance, on the other hand, has to do with how courts go about performing their roles and whether their operations conform to various standards of proper judicial practice, fairness, and efficiency. To assess process performance is thus to look at judicial procedures. Relevant procedures include how courts receive claims and how they handle claims once received. They also include how courts process evidence, deliberate, arrive at decisions, and develop arguments as well as how they interact with disputants and other interested parties. Simply put, the analysis of process performance encompasses the evaluation of the full range of judicial procedures.⁹ Our thinking about process performance draws on the work of Gutner

⁸ In practice, the Council has not exercised this authority.

⁹ We view judicial appointment processes as external to the judicial proceedings, and thus do not include it per se in process performance. Nevertheless, judicial appointment processes may be instrumental in shaping the performance of courts.

and Thompson (2010). But we differ from the approach of Shany (2014) who treats judicial procedures primarily as a determinant of goal attainment.

Criteria for evaluating process performance rest on normative preferences regarding judicial procedures. These preferences may or may not be specified by the framers of an IC. If we posit a preference for processes that are fair, for example, we can evaluate the extent to which an international court's procedures meet major standards of procedural fairness. Broadly speaking, the standards we apply in assessing the process performance of ICs follow rule of law principles familiar from the study of domestic courts. Such principles accentuate matters of procedural fairness or procedural justice. Additional criteria relevant to the evaluation of process performance include matters of efficiency covering both the time it takes to render judgments on specific cases and the material costs of operating a judicial system. In cases where caseloads are heavy and resources are limited, the ability of a court to handle individual cases with minimal expenditure of time and resources becomes particularly important. Other criteria for process performance may include standards relating to methods and procedures of judicial decision-making (e.g., transparency, reasoning and interpretation, fact-finding).

Outcome performance and process performance are distinct. In principle, at least, a court may perform well in settling disputes or encouraging compliance, even when its performance with regard to process standards is relatively poor. Nevertheless, many analysts have taken an interest in the relationship between the two, exploring whether performing well in terms of process contributes to outcome performance. The issue here resembles the debate about the extent to which “good” governance contributes to effective governance. Process performance may be valued as an end in itself, but many value process performance on instrumental grounds on the assumption that performing well in process terms contributes positively to performing well in outcome terms. In our view, this is a matter worthy of systematic analysis among those interested in the performance of international courts.

Table 1.1 summarizes this discussion of criteria for assessing the performance of ICs and provides a template for the substantive chapters in this volume.

Table 1.1 The Performance of ICs

<i>Performance Dimension</i>		
<i>Performance criteria</i>	<u>Process</u>	<u>Outcome</u>
	<ul style="list-style-type: none"> • Procedural fairness • Efficiency standards • Standards of legal reasoning and interpretation 	<ul style="list-style-type: none"> • Dispute settlement • Clarification of law • Compliance

Levels of Analysis

Implicit in our discussion so far is a focus on the individual case or judicial action. Strictly speaking, what courts do is to accept or decline to intervene in the claims of those seeking judicial intervention and to render judgments on those cases that make it onto their dockets. But it is relevant to think about the consequences of the judgments of international courts at other levels as well. In this regard, we distinguish among the micro-, meso-, and macro-levels, corresponding roughly to the level of the individual case, the level of the issue areas or governance arrangements to which cases belong, and the level of the overarching system in which cases arise. In simple terms, the meso- and macro-levels concern how courts perform in relation to broader outcomes or processes, as opposed to outcomes or processes at the level of a specific dispute or legal case.

Like all other courts, ICs render judgments regarding specific cases arising within their sphere of jurisdiction. Individual cases may take the form of a dispute between two or more states, a dispute between one or more state and a non-state actor, or an evaluation of allegations that an individual’s behavior constitutes criminal actions. Performance at this level, which we call micro-level performance, constitutes the initial concern in any assessment of the

performance of international courts. Though courts may seek to arrive at decisions on the narrowest possible grounds to minimize the consequences or impacts of their actions beyond the cases at hand, this is not the only level of analysis at which we can assess the performance of courts. In the course of handing down judgments, the actions of courts often produce effects extending well beyond the level of the individual cases. Thus, we can look to IC performance beyond specific cases to examine how ICs produce effects involving broader outcomes and processes.

When the judgments of ICs affect an issue area or governance system, we characterize the results as meso-level performance. Many regional integration regimes, for example, seek to promote economic and political integration among their members. In connection with these regimes, we can ask whether and how the judgments of a related international court have contributed to the pursuit of this goal. The CJEU provides a well-documented instance of an international court playing a vital role in advancing economic and political integration (Burley and Mattli 1993; Helfer and Slaughter 1997; Mattli and Slaughter 1998; Stone Sweet and Brunell 1998). Similarly, regional human rights regimes often seek to strengthen human rights law and policy. Some analyses of the performance of the ECtHR adopt this perspective, asking how the Court has contributed to the achievement of this goal (Helfer and Voeten 2014). Another way to think about performance at the meso-level is to ask whether the activities of a court have contributed to the prevention of disputes and behavioral change across the relevant governance system. International courts can contribute to dispute prevention by reducing uncertainty about the precise meaning of rules and encouraging parties to settle disputes before adjudication arises (for example, Busch and Reinhardt 2006; Kim 2008; Gilligan et al. 2010). Similarly, parties may adjust their behavior *ex ante* to avoid incurring costs associated with judicial oversight. An IC thus may perform well if it reduces the number of cases brought before it in the long term. A central debate concerning the performance of international criminal courts

and tribunals, for example, centers on the question of whether they play a role in deterring war crimes going forward (Wippman 1999; Akhavan 2001; Gilligan 2006; Rodman 2008; Kim and Sikkink 2010).

A third level of performance, which we call macro-level performance, directs attention to the impacts of an international court's activity on international society more generally. Of course, the higher the level of analysis, the more abstract the criteria for evaluating performance become and the longer the causal chains involved. Nevertheless, it is worth asking whether the activities of international courts (individually or collectively) make a difference in terms of the character or state of international society (Alter 2014). Thus, some observers have asked whether the activities of the WTO's dispute panels or AB have empowered developing nations. Similarly, the activities of courts in some settings have contributed to judicialization in other settings. There is evidence to suggest that the activities of ICTY and the ICTR influenced the international community's decision to establish the ICC. Similarly, the CJEU has served as a model for those seeking to create judicial mechanisms to handle disputes and other legal problems in a number of other regions (Alter 2012).

For the most part, the discussion of levels of performance addresses what we have called outcome performance. But it is worth asking whether there are similar distinctions with regard to process performance. Process performance at the micro-level centers on matters like procedural fairness and efficiency. These process considerations remain important at the meso-level, but we are also concerned with the development of a coherent or internally consistent body of law. At the macro-level, we can ask whether international courts are playing a role in spreading and strengthening the influence of the collection of practices generally described in terms of the rule of law. Although the primary reason to differentiate among levels involves outcome performance, then, there is a parallel set of distinctions relating to process performance that are of interest in the study of the performance of international courts.

Before taking the next step, we want to draw attention to the fact that there are challenges to explaining variations in performance across ICs or over time with regard to the same court. The central point is that links between outcomes of interest and the actions of courts are often difficult to establish precisely or conclusively. Sometimes the causal chain is short and the causal mechanisms are simple. If we ask questions about the decisions of courts regarding the acceptance of cases for consideration or about the contents of the judgment they render with regard to a specific case, there is no doubt about the causal connections. Courts are authors of their own decisions. But as soon as we proceed beyond these simple links, the causal connections become more complex. If the parties to a dispute choose to adjust their behavior to comply with the terms of a judicial judgment, for example, can we attribute their actions solely to the authority of the court? It does not take much analysis to realize that the decisions of parties regarding compliance are influenced by a range of factors that extend beyond the actions of the relevant courts. Establishing causal connections between the activities of courts and outcomes of interest becomes progressively more difficult as we move from the micro-level to the macro-level because causality becomes more complex and causal chains become longer. It is hard enough to establish causal connections at the micro-level. Proving that courts make a difference beyond the level of specific cases in issue domains like human rights or trade requires the use of clever methodological strategies. Demonstrating that the actions of courts influence the nature of international society as a whole is still more challenging. Despite these difficulties, it would be a mistake to conclude that courts are unimportant at the meso- and macro-levels. But it is important to exercise caution in arriving at judgments regarding the performance of courts at these levels. Still, we take the view that it is worthwhile to consider these higher levels of performance to gain a comprehensive understanding of the performance of ICs and to assess the significance of arguments regarding the legalization in international society.

Determinants of Performance

Obviously, tracking the performance of international courts is a demanding task. But assume for the moment that we have a clear understanding of the concept of performance and some ability empirically to measure or track patterns and variations in the performance of the ICs of interest to this project. The next challenge is to seek explanations of variations and patterns in performance we are able to document. We want to know why some courts do better than others in terms of performance, whether there are significant variations in performance among courts operating in different issue areas or in different regions of the world, and why the performance of the same court may rise and fall over the course of time. This leads to a search for the determinants of performance. In more formal terms, if we think of performance as the dependent variable (or perhaps as a suite of dependent variables) in this context, we now want to direct attention to the search for independent variables that allow us to explain or account for variance in the dependent variable.

Several introductory comments are in order as we embark on this journey. To begin with, we want to evaluate performance in terms of two dimensions (process and outcome performance), three levels (micro-, meso, and macro-level performance), and four issue areas (trade, investment, human rights, and criminal actions). There is no reason to assume a priori that the performance of courts will be uniform or formative across dimensions, levels or issues. High performance with regard to the dimension of process may not translate into high performance with regard to outcomes. High performance at the micro-level need not lead to high performance at the meso-level, much less the macro-level. The fact that courts perform well in one issue area (e.g. investment) tells us little about how well they will perform in other issue areas. Our initial expectation is that we are likely to encounter a great deal of variance in examining the performance of international courts across the dimensions, levels and issue areas.

In all cases, moreover, we must come to grips with issues of causal inference. It is helpful, as a point of departure, to identify associations or correlations between variations we are able to document regarding performance and factors that we hypothesize to be determinants of performance. But how can we be sure that the relationships we are examining are not spurious? If we observe that disputants accept the judgments of courts more readily in social settings in which power is highly concentrated, for instance, can we conclude that the concentration of power is the key determinant? Similarly, if courts have more success in clarifying the law in settings characterized by a high level of consensus regarding principles of judicial interpretation, are we justified in treating consensus as the essential determinant of success? Causal inference can also be complicated by complex causality. In some cases, the causal chain is short and comparatively simple. This is particularly true at the micro-level where we are concerned with the ability of courts to render judgments regarding the cases that come before them, the likelihood that the parties will accept or comply with the judgments that courts hand down, and the prospect that the handling of cases will make a difference regarding the treatment of the problems at hand. But even here it is important to be cautious in drawing causal inferences about factors that determine the performance of courts. Many factors may come into play in determining whether parties comply with the judgments of a court, for example. And when we turn to the influence of the actions of courts at the meso- and macro-levels, it is clear that we are operating in the realm of complex causality. There are no magic solutions to the problem of demonstrating causality in this realm; all the contributors to this volume wrestle with this challenge in arriving at conclusions pertaining to their individual topics.

This is not to diminish, much less dismiss, research that traces the role of courts in international society. But it does mean that we will typically find ourselves dealing with issues featuring complex causality in the sense that the actions of courts are significant as elements of what we can think of as causal clusters. In some instances, we will want to make an effort to

devise means to tease out the contributions of ICs or, in more formal terms, to determine the proportion of the variance in outcomes of interest attributable to the activities of courts. In other cases, however, it may be sufficient to consider the influence of clusters of determinants, without making a sustained effort to separate out the influence of the activities of courts from a range of other factors. As we move toward the meso- and macro-levels, a focus on the character of such clusters seems likely to become increasingly important.

That said, we can identify some of the principal lines of argument that analysts bring to bear in efforts to explain the performance of courts. Our aim in this regard is not to spell out a set of precise hypotheses that we expect authors of substantive chapters to examine or “test” systematically in their contributions. Rather, we seek to identify and comment on the main types of explanations that those who think about such matters have articulated. Our expectation is that these observations will play a role in guiding the work of the substantive chapters in this volume, as well as future research, and that we can come back to in the conclusion to reflect on the relative importance of various factors in accounting for variations in the performance of international courts.

One way to think about determinants of the performance of courts is to turn to the main streams of analysis embedded in contemporary work on international relations more broadly: realism, liberalism, institutionalism, and constructivism. Realists focus on power and particularly on the configuration of power in specific settings in the search for explanations. Realists who study international law and courts place central importance on state interests and the distribution of power among states (Goldsmith and Posner 2005; Posner and Yoo 2005). This mode of analysis yields two main explanatory perspectives regarding the role of ICs. Is there a dominant state (that is, what the realists call a hegemon) or a dominant coalition of states that provides backing for a court and supports its role in a given issue area? The idea here is that the performance of courts reflects in a general way the power structure prevailing in a given

social setting. In this regard, we would expect the performance of courts to suffer in cases where the power structure is contested or in flux. Realists also look to matters of problem structure and ask whether the issues that give rise to cases that come before courts amount to coordination or collaboration problems (Goldsmith and Posner 2005). Similarly, realists expect better performance regarding issues involving low versus high politics (Zangl et al. 2012). While courts may do well on this account in handling coordination problems or low politics issues, their performance is expected to be poor if and when they tackle collaboration problems or issues of high politics.

Liberals and institutionalists, by contrast, think of the role of ICs as organizations that facilitate the operation of institutions in specific issue areas. This may be a matter of resolving ambiguities in the application of regulatory arrangements or resolving disagreements about the extent to which actors are complying with regulatory arrangements. In other cases, it will be a matter of sorting out situations in which there is disagreement regarding the requirements of regulatory measures associated with different regimes (e.g. trade regimes and environmental regimes). ICs on this account may be either general arrangements (e.g. the ICJ) whose jurisdiction allows them to address issues arising in many issue areas or more specialized arrangements (e.g. the ECtHR) whose jurisdiction extends to cases arising in a well-defined issue area. In the liberal view, domestic actors and institutions influence the performance of international courts (Helfer and Slaughter 1997; Mattli and Slaughter 1998; Alter 2001; Davis 2012). For example, performance may hinge on whether relevant state parties are democratic, the type of legal tradition, domestic courts and domestic society. Institutionalism, on the other hand, views performance as a function of institutional design of ICs (Helfer and Slaughter 1997; Keohane et al. 2000; Helfer and Slaughter 2005; Zangl 2008; Zangl et al. 2012). Those who adopt this perspective deem access for private actors, independence, and enforcement capacity as well as the precision and obligation of international law to be instrumental to performance.

For their part, constructivists associate the role of ICs with systems of ideas embedded in society. A particularly powerful example is the idea of the rule of law with its presumption that disputants should resort to legal procedures to resolve their differences rather than falling back on the use of force and that legal proceedings constitute the appropriate means for determining whether or not parties are in compliance with prevailing rules in specific cases. A particularly important application of constructivism involves the growth of new rights and rules that may take precedence over the presumption of sovereign immunity. The creation and operation of international courts dealing with human rights, for example, is a function of the growth in recent decades of principles and rules relating to human rights that states are expected to adhere to regardless of the doctrine of sovereign immunity. The current effort to develop and make use of international laws dealing with criminal behavior is a particularly interesting example of the growth of a new international discourse governing the actions of states as well as individuals. The extent to which the International Criminal Court thrives in the coming years will be an important indicator of the growth of such a discourse on a global scale. In all cases, constructivists look to systems of ideas and the extent to which they are pervasive in international society for explanations of variance in the performance of ICs. For this reason, constructivists consider processes of socialization and norm internalization to be critical to understanding patterns of performance (Lutz and Sikkink 2001; Sikkink 2011; Goodman and Jinks 2013).

Another way to think about determinants of the performance of international courts is to compare and contrast the perspectives of political scientists and lawyers regarding factors that can account for variation in performance. Of course, such differences are not absolute. But political scientists often point to the importance of judicial politics as determinants of the performance of courts. The scope of this concern involves both intra-court dynamics and relations between courts and those who make the relevant rules and those who are responsible

for implementing the judgments of courts in specific situations. Internal battles are common within courts, and it is reasonable to expect that the opinions of (near) unanimous courts will carry more weight when it comes to various measures of performance than those of deeply divided courts. Among other things, the opinions of deeply divided courts invite those on the losing side to try again with slightly different cases in the hope of getting a different outcome in new cases. Beyond this, courts may enjoy more or less comfortable relations with those responsible for making and implementing laws. Political scientists debate the extent to which performance hinges on how well aligned a court's decisions are with the preferences of states (Garrett et al. 1998; Cichowski 2007; Alter 2008; Carrubba et al. 2008; Stone Sweet and Brunell 2012; Stone Sweet and Brunell 2013). Embattled courts may suffer with regard to performance, even though they are highly visible players in addressing some social issues. On the other hand, courts that get along well with those who make and implement the rules may have little influence at the meso- and macro-levels. Another approach taken by some political scientists concerns the influence of various types of nonstate and substate actors. Strong performance of courts may or may not depend upon higher degrees of public support or perceived legitimacy (Gibson and Caldeira 1995; Gibson and Caldeira 1998). Similarly, performance may be associated with interactions with "compliance partners," such as national courts, domestic constituencies, and activists (Alter 2001; Dai 2005; Cichowski 2007; Simmons 2009; Alter 2014;).

Lawyers, by contrast, often highlight the importance of factors such as jurisdiction, admissibility, and independence and impartiality (Helfer and Slaughter 1997; Shany 2014). Lawyers tend to pay greater attention to methods of judicial interpretation. They often differentiate between activist courts that endeavor to play an influential role regarding major social issues (e.g. the enlargement of civil and human rights) and more passive courts that seek to resolve the issues that come before them on narrow grounds. Similarly, there are differences

between those who regard themselves as strict constructionists and seek to ground their opinions in the original meaning of prevailing rules and those who are willing to construe laws in such a way as to adapt them to changing circumstances or to advance social goals they espouse. The underlying idea in this connection is that matters pertaining to methods of interpretation may explain the performance of ICs in specific settings, though there is no reason to assume that one method or another will always strengthen the performance of courts.

Beyond this are a number of other perspectives on the determinants of performance that do not fit neatly into the categories outlined in the preceding paragraphs. While it is possible to devise arguments about the potential relevance of a wide range of factors, three additional perspectives have emerged in our preliminary discussions that deserve separate comment in connection with the thinking embedded in this project: (i) types of cases, (ii) the normative pull of judicial actions, and (iii) the role of leadership.

Traditional international law recognizes only states as subjects, an arrangement that limits litigation to cases involving two or more states willing to accept the jurisdiction of a court. But increasingly, states must accept the compulsory jurisdiction of a court if they want to become part of an international regime, and many international courts today consider a wider range of cases, including those in which private actors bring complaints against public authorities and those in which public authorities levy charges against individuals alleged to have engaged in criminal activities. This makes it possible to ask whether the performance of courts differs from one type of case to another. Cases in which both parties are states are notoriously difficult to deal with due to the barriers imposed by sovereignty and the ability of states to refuse to accept the jurisdiction of courts with regard to issues they do not intend to subject to the rule of law. It is interesting to ask, under the circumstances, whether we can expect international courts to perform better when they are dealing with cases involving non-state actors.

Another line of thinking that is worthy of consideration centers on the issue of what some observers characterize as the “normative pull” of legal judgments. There is clearly a relationship between the normative principles embedded in the judgments of courts and the evolution of social discourses in any given societal setting. When courts get too far ahead of the state of the social discourse in an area like standards applicable to human rights, their decisions are apt to be ignored by most members of international society. Conversely, there are cases when social discourse evolves quickly, and the role of courts is to adapt legal standards to bring them into line with broader societal developments. In most cases, however, the decisions of courts are likely to be stimulants in the progressive development of international norms. The idea here is that the actions of courts exercise a normative pull that provides actors with an incentive to adapt existing norms or adopt new ones and adjust their behavior accordingly. We can ask, in this connection, whether there is an optimal balance between the prevailing practices and the normative content of court decisions as a determinant of the performance of international courts. Considerations of this sort are particularly relevant to efforts to assess the performance of ICs at what we have described as the meso-level.

Although leadership is discussed more often in the context of executive actions or legislative politics than in connection with the activities of courts, there is no doubt that judicial leadership can play a role as a determinant of the performance of international courts. Two distinct perspectives on leadership are notable in this context. One involves the ability of individual judges to forge consensus among the members of a court regarding both the content of judgments rendered in specific cases and the legal arguments crafted in support of these judgments. The other involves the ability of judges to articulate legal principles that gain credence at the societal level and emerge as influential determinants of legal and political practices beyond the confines of the case at hand. As those interested in the role of courts have observed regularly, the dicta articulated in legal opinions (even in cases where they are set forth

in minority or dissenting opinions) can become influential going forward when they express normative principles in a concise and appealing manner or when they articulate sentiments that are nascent in the social discourses of a particular time and place. Here, again, it is interesting to think about the effects of the actions of courts at what we have characterized as the meso-level in contrast to the micro-level. A court does not have to compile a perfect record with regard to its performance at the micro-level in order to make a difference at the meso-level. A line of decisions that move the social discourse associated with a particular issue area in some well-defined direction can produce a significant effect at the meso-level that extends well beyond the disposition of individual cases.

As the observations in the preceding paragraphs suggest, there may be significant differences in the determinants of the performance of ICs across the levels of analysis we have identified. Of course, it is difficult to imagine that a court could make a real difference at the meso-level, much less the macro-level, if it performs poorly at the micro-level. But some courts might make a conscious effort to render decisions regarding specific cases on narrow grounds to minimize any impact their decisions have at the meso- and macro-levels. Conversely, the broader impacts of the actions of courts may survive setbacks with regard to the disposition of individual cases. What this suggests is that it is useful to draw clear distinctions among the three levels of analysis in examining the determinants of the performance of international courts. Causal chains become longer and more complex as we move from an assessment of the performance of courts at the micro-level to their performance at the meso- and especially the macro-levels. This means our judgments will become less robust as we move from one level to another, but this does not mean that ICs do not play influential roles with regard to the evolution of major issue areas or the character of international society as a whole. So long as we are careful to qualify our assessments appropriately, this seems to us to be a worthwhile line of analysis.

The shape of things to come

The substantive chapters in this book comprise two major clusters of contributions followed by a conclusion that seeks to articulate preliminary responses to the questions we have formulated in this introductory chapter and to suggest fruitful avenues for further analysis.

Part I consists of a set of chapters providing broad assessments of the role of international courts in a number of distinct issue areas, including trade, investment, human rights, and international criminal law. Each of these chapters seeks to collect and synthesize existing research in the relevant issue area rather than to report the findings of new research. The authors of these chapters apply the performance criteria that we have identified in this chapter. But as we have argued, not all criteria equally apply to all ICs nor have we been exhaustive. We provide a general template that authors adapt to fit the issue area they discuss. Thus, the chapters in Part I apply the criteria that are most fitting for the issue they examine while also bearing in mind the state of existing research and debates of that issue area.

Part I begins with two chapters focused on adjudication of international economic law. In Chapter Two, Cosette Creamer and Anton Strezhnev assess the performance of ICs within the multilateral trade regime, focusing on inter-state disputes and the WTO Dispute Settlement Mechanism. Drawing on existing literature and original data, Creamer and Strezhnev argue that the WTO DSM is most often able to facilitate the resolution of conflicts relating to market access. While there may be some question as to what “resolved” means, most WTO disputes are resolved in a manner acceptable to all parties and resulting in high levels of compliance. Less clear, however, is whether the WTO DSM encourages pre-litigation or post-litigation resolution. The authors argue that the dispute settlement mechanism has an informal role in clarifying international trade law, which states generally accept. Creamer and Strezhnev argue

the WTO DSM process performance is weakened by its inability to ensure access to justice. Overrepresentation of the most powerful economies is problematic in the WTO DSM because disparities in legal capacity hinder equal access. The overrepresentation of powerful economies in the DSM also has implications for outcome performance, especially in terms of the development of law. Last, their chapter reviews other dispute settlement processes embedded within preferential trade agreements, such as the CJEU.

In Chapter Three, Daniel Behn evaluates the performance of international investment treaty arbitration. He argues that investment treaty arbitration has contributed to legal consistency emerging on legal rules in international investment law. Behn shows that the performance of international investment arbitration suffers from challenges of access as investment agreements do not offer universal coverage to investors and a disproportionate amount of litigation arises from only a portion of international investment agreements. In terms of outcomes, investors and respondent states tend to win at equal rates on the merits. But when considering jurisdictional decisions, discontinuances and negotiated settlements, Behn shows that there is a pro-developed state bias in investment arbitration. As for compliance, most arbitral judgments are enforced, even though there are notable instances of state defiance. The process performance of investment treaty arbitration is mixed. Behn argues that while there is increasing transparency in the regime, the regime's performance suffers from lengthy resolution of disputes and a lack of diversity among the arbitrators appointed to sit on tribunals.

Turning to the performance of regional human rights courts, Chapter Four by Dinah Shelton evaluates the performance of the ECtHR, the IACtHR, and the African Court of Human and Peoples' Rights (AfCHPR). Shelton discusses how internal structural factors, such as tenure and appointment of judges, and external structural factors, like a lack of resources, affect the performance of regional human rights courts. She argues that the regional human rights courts perform better than member states expect but not as well as victims would hope. Their process

performance suffers from limited access, lengthy proceedings, and the courts' inability to address widespread and serious violations. The outcome performance of the regional human rights courts is mixed due to variation in the degree of deference granted to states and the forms of redress provided to victims. Nevertheless, she argues all three courts generally perform well in terms of generating compliance with their decisions, even though compliance with the IACtHR is lower than its European counterpart. In addition, they have contributed to the development of normative principles relating to human rights.

Chapter Five by Nobuo Hayashi examines the outcome performance of international criminal courts and tribunals (ICTs). Of interest here is one particularly controversial aspect of ICTs' outcome performance: whether they contribute to the development of international humanitarian law (IHL) and international criminal law (ICL). Non-retroactivity and strict interpretation, two cardinal principles of criminal justice, make reliance on custom as a source of punishable violations of IHL awkward. Yet, ICTs have engaged routinely in what some regard as interstitial law-making by their judges. This chapter argues that custom as a source of ICL will endure despite the Rome Statute's relatively high degree of codification, and that ICT judges will continue to claim to find customary rules even where the evidence for them is thin or non-existent. It will therefore be crucial to see how states, the pre-eminent makers of today's customary international law, will respond to such claims. The chapter reveals that, while some forward-looking pronouncements find subsequent support amongst states and become codified, others elicit very different reactions. State involvement in ICT adjudication can sometimes make a difference, but not always. Hayashi concludes that overall the normative contribution of ICTs to IHL and ICL has been more mixed than is often claimed.

The chapters in Part II then turn to efforts to identify factors relevant to understanding the performance of courts. Rather than revisiting familiar themes, these chapters seek to illuminate important determinants of IC performance that have eluded rigorous assessment in

prior work on ICs. The point is not to argue that familiar themes (e.g., the court's jurisdiction, judicial appointment procedures, legitimacy, etc.) are unimportant but rather to add to an understanding of factors that have not been highlighted in previous studies of the performance of ICs. These chapters feature original research that applies the concept of performance to concrete instances of IC performance. Each of chapter contributes to our understanding of whether, how and why IOs perform as they do. The chapters in Part II are methodologically and empirically diverse. While some chapters dig into the performance of a single IC in a single case, others compare across several ICs. Some chapters examine process performance and other outcome performance. The chapters feature statistical data analysis as well as qualitative data analysis. The richness is the variation of these chapters illustrates how the concept of performance can be empirically applied across the dimensions, levels and issues areas pertaining to international law and courts.

In Chapter Six, Hyeran Jo, Mitchell Radtke, and Beth Simmons analyze the ability of the ICC to deter. Focusing on outcome performance, they argue that the ICC's performance should be assessed in terms of its ability to deter crimes that are subject to its jurisdiction. The chapter presents two kinds of evidence to assess the outcome performance of the ICC: first, whether it has had an imprint on domestic law, and second, whether there is evidence that the ICC or the domestic laws it has encouraged have had deterrent effects. The chapter relies on broad-based evidence to evaluate the ICC's deterrence effects, while also considering in depth the case of Uganda. Using a time-series intervention analysis, the authors provide robust evidence that violence against civilians by state agents and rebels dropped off significantly following the ICC's intervention in Uganda. While it is not conclusive, this finding constitutes significant evidence that the ICC has performed well against one exceedingly important criterion: deterring the killing of civilians in the context of civil wars.

Chapter Seven by Benjamin Faude explores the link between institutional fragmentation and IC performance. Faude considers how the proliferation of ICs challenges the commonly held assumption that ICs do not interact with one another. His central argument is that forum shopping may occur as a result of institutional fragmentation and that this can generate a return to politicized resolution of inter-state conflicts. Divergent decisions from ICs with overlapping jurisdiction may compromise the performance of ICs and lead parties to solve their disputes outside of international judicial bodies. However, if ICs take each other's case law into account, they may be able to boost each other's performance, effectively foreclosing forum-shopping opportunities. Judicial dialogue and informal recognition of the judgments of other ICs as precedential may lead to the clear allocation of competencies among overlapping ICs or to the reconciliation of their jurisprudence. Faude looks at these dynamics in disputes involving softwood lumber brought before the DSMs of the North American Free Trade Agreement and the WTO.

Jeff Dunoff and Mark Pollack focus in Chapter Eight on process performance. The authors broaden the theoretical and methodological approaches used in evaluating court performance by introducing "practice theory" to evaluate judicial performance. They argue that to understand the performance of international courts it is necessary to assess the performances of judges. Judicial performances are based not only on court statutes and rules, but also are created by judges themselves in the practices they adopt. Dunoff and Pollack explain that judicial practices exist throughout a dispute's lifecycle and include multiple performances of judges, ranging from their private deliberations with colleagues to their off-the-bench activities such as delivering speeches. The chapter highlights the evidentiary and fact-finding practices of the International Court of Justice, offering contrasts to other ICs, and links these practices to the outcome performance of the Court.

In Chapter Nine, Nicole De Silva explores how ICs try to influence their own performance by focusing on strategies of socialization. De Silva argues that ICs can formulate policies and practices aimed directly at socializing actors into the norms, rules, and procedures that underpin the performance of ICs. In these strategies, ICs move beyond adjudication and their mandated activities, developing policies and practices that feature diplomacy, training, and outreach, for socializing relevant actors in their legal regimes. The chapter presents a framework for conceptualizing ICs' socialization strategies and their influence on the actual and perceived performance of ICs. De Silva shows how ICs can formulate socialization strategies in response to challenges to their actual and perceived performance, thus highlighting a realm of IC decision-making and activity that has received little attention in existing scholarship. The chapter's analysis of ICs' reported policies and practices shows that these socialization strategies are an important means by which ICs aim to influence their performance and actors' perceptions of it.

Chapter Ten by Chiara Giorgetti concentrates on factors relating to compliance. Specifically, Giorgetti reviews the compliance mechanisms international courts use and discusses political, sociological, and other factors that affect compliance. She argues that compliance with court judgments is significant in determining whether a court is effective and performs as expected. She applies this line of analysis to four ICs: the ICJ, ECtHR, tribunals under the auspices of the International Convention for the Settlement of Investment Disputes, and the United Nations Claims Commission. The aim is to offer an appraisal of a variety of mechanisms that mirror the full panoply of existing compliance mechanisms as well as to evaluate their operability in 'real life.' Giorgetti argues that these formal mechanisms of compliance contribute to judgment compliance to varying degrees. Several legal factors affect compliance, including the types of remedies involved, the cost of compliance, the kind of case, and the composition of the court. Giorgetti also explains how external political pressure and

features of the state, such as its domestic judicial system, contribute to compliance with IC judgments.

In Chapter Eleven, Steinar Andresen asks what the analysis of IC performance can learn from the literature on international regimes. Andresen argues that there is much to be learned from this literature, especially for understanding the meso-level performance of ICs. This chapter discusses how regime effectiveness literature reveals particular methodological challenges in assessments of ICs' performance. Specifically, while we may be able to attribute causal significance to an IC, determining how much an IC matters is difficult. He encourages scholars to view ICs as intermediate variables, in the sense that they channel and guide human interaction in the issue areas in which they operate. Additionally, he argues that the regime literature suggests that it is pertinent to focus on problem structure, asking to what extent ICs are able to deal effectively with difficult or "malign" issues. He also points to other explanatory factors worth considering, especially the broader context, including adherence to "the rule of law" and the problem-solving capacity of ICs. Commenting on the impact of courts in the area of human rights, Andresen suggests the performance of regional human rights courts is linked to how long they have existed, their participatory scope, and the severity of the challenges facing them.

Part III summarizes key findings and discusses future directions for research on the performance of international courts. Chapter Twelve by Theresa Squatrito focuses on methodological considerations for future research on IC performance. She discusses the operationalization and measurement of performance, reviewing existing and possible approaches to measurement and their advantages and disadvantages. In addition, she canvases some options for methods that researchers might use to evaluate and explain IC performance. In particular, she discusses five methods—experimental designs, meta-analysis, qualitative

comparative analysis, counterfactual analysis, and mixed methods research—and how they might be useful for future research on the performance of international courts.

Chapter Thirteen returns to our basic questions about the performance of ICs. How well have ICs performed, and does performance vary by issue area? What are the determinants of performance in this realm? What can we expect regarding trends in performance during the foreseeable future? We note that much of the work carried out in this field so far highlights what we have described as outcome performance at the micro-level. Even in this realm, our contributors are on more solid ground in describing performance than in explaining the patterns they identify. This is understandable given the challenges of arriving at convincing causal judgments regarding this category of performance, much less process performance in contrast to outcome performance and performance at other levels. This means that there is a broad range of issues relating to the performance of ICs that remain to be tackled in a sustained manner by analysts seeking to assess the judicialization of international affairs and more specifically the roles that ICs play in this development. While progress in addressing these issues will require a combination of ingenuity and hard work, it is hard to exaggerate the importance of deepening our understanding of this subject.

References

- Akhavan, Payam 2001. "Beyond Impunity: Can International Criminal Justice Prevent Future Atrocities?," *American Journal of International Law* 95(1): 7-31.
- Alter, Karen J. 2001. *Establishing the Supremacy of European Law: The Making of an International Rule of Law in Europe*. Oxford: Oxford University Press.
- 2008. "Agents or Trustees? International Courts in Their Political Context," *European Journal of International Relations* 14(1): 33-63.
- 2012. "The Global Spread of European Style International Courts," *West European Politics* 35(1): 135-154.
- 2013. "The Multiple Roles of International Courts and Tribunals: Enforcement, Dispute Settlement, Constitutional and Administrative Review," in Dunoff, Jeffrey and Pollack, Mark (eds.), *Interdisciplinary Perspectives on International Law and International Relations*. Cambridge: Cambridge University Press, pp. 345-370.
- 2014. *The New Terrain of International Law: Courts, Politics, Rights*. Princeton: Princeton University Press.
- Alter, Karen J. and Helfer, Laurence R. 2010. "Nature or Nurture? Judicial Lawmaking in the European Court of Justice and the Andean Tribunal of Justice," *International Organization* 64(4): 563-592.
- Alvarez, José 2014. "What Are International Judges For? The Main Functions of International Adjudication," in Romano, Cesare P.R., Alter, Karen J. and Shany, Yuval (eds.), *The Oxford Handbook of International Adjudication*. Oxford: Oxford University Press, pp. 159-178.
- Bogdandy, Armin Von and Venzke, Ingo 2013. "On the Functions of International Courts: An Appraisal in Light of Their Burgeoning Public Authority," *Leiden Journal of International Law* 26(1): 49-72.
- Breitmeier, Helmut, Young, Oran R. and Zürn, Michael 2006. *Analyzing International Environmental Regimes*. Cambridge, MA: MIT Press.
- Buchanan, Allen and Keohane, Robert 2006. "The Legitimacy of Global Governance Institutions," *Ethics and International Affairs* 20(4): 405-437.
- Burley, Anne-Marie and Mattli, Walter 1993. "Europe before the Court: A Political Theory of Legal Integration," *International Organization* 47(1): 41-76.
- Busch, Marc L. and Reinhardt, Eric 2006. "Three's a Crowd: Third Parties and WTO Dispute Settlement," *World Politics* 58(3): 446-477.
- Caron, David C. 2006. "Towards a Political Theory of International Courts and Tribunals," *Berkeley Journal of International Law* 24(2): 401-422.
- Carrubba, Clifford J., Gabel, Matthew and Hankla, Charles 2008. "Judicial Behavior under Political Constraints: Evidence from the European Court of Justice," *American Political Science Review* 102(4): 435-452.
- Cichowski, Rachel A. 2007. *The European Court and Civil Society: Litigation, Mobilization and Governance*. Cambridge: Cambridge University Press.
- Dai, Xinyuan 2005. "Why Comply? The Domestic Constituency Mechanisms," *International Organization* 59(2): 363-398.
- Davis, Christina 2012. *Why Adjudicate?: Enforcing Trade Rules in the WTO*. Princeton: Princeton University Press.
- Garrett, Geoffrey, Kelemen, R. Daniel and Schulz, Heiner 1998. "The European Court of Justice, National Governments, and Legal Integration in the European Union," *International Organization* 52(1): 149-176.

- Gibson, James L. and Caldeira, Gregory A. 1995. "The Legitimacy of Transnational Legal Institutions: Compliance, Support, and the European Court of Justice," *American Journal of Political Science* 39(2): 459-489.
- 1998. "Changes in the Legitimacy of the European Court of Justice: A Post-Maastricht Analysis," *British Journal of Political Science* 28(1): 63-91.
- Gilligan, Michael J. 2006. "Is Enforcement Necessary for Effectiveness? A Model of the International Criminal Regime," *International Organization* 60(4): 935-967.
- Gilligan, Michael, Rosendorff, B. Peter and Johns, Leslie 2010. "Strengthening International Courts and the Early Settlement of Disputes," *Journal of Conflict Resolution* 54(1): 5-38.
- Goldsmith, Jack and Posner, Eric A. 2005. *The Limits of International Law*. Oxford: Oxford University Press.
- Goldstein, Judith, Kahler, Miles, Keohane, Robert O. and Slaughter, Anne-Marie 2000. "Introduction: Legalization and World Politics," *International Organization* 54(3): 385-399.
- Goldstein, Judith L., Rivers, Douglas and Tomz, Michael 2007. "Institutions in International Relations: Understanding the Effects of the Gatt and the WTO on World Trade," *International Organization* 61(1): 37-67.
- Goodman, Ryan and Jinks, Derek 2013. *Socializing States Promoting Human Rights through International Law*. Oxford: Oxford University Press.
- Gutner, Tamar and Thompson, Alexander 2010. "The Politics of IO Performance: A Framework," *Review of International Organizations* 5(3): 227-248.
- Guzman, Andrew T. 2008. "International Tribunals: A Rational Choice Analysis," *University of Pennsylvania Law Review* 157(1): 171-236.
- Hafner-Burton, Emilie M. 2013. *Making Human Rights a Reality*. Princeton: Princeton University Press.
- Haye, Eve 2008. *War Crimes in Internal Armed Conflicts*. Cambridge: Cambridge University Press.
- Helfer, Laurence R. 2014. "The Effectiveness of International Adjudicators," in Romano, Cesare P.R., Alter, Karen J. and Shany, Yuval (eds.), *The Oxford Handbook of International Adjudication*. Oxford: Oxford University Press, pp. 464-482.
- Helfer, Laurence R. and Alter, Karen J. 2013. "Legitimacy and Lawmaking: A Tale of Three International Courts," *Theoretical Inquiries in Law* 14(2): 479-503.
- Helfer, Laurence R., Alter, Karen J. and Guertzovich, M. Florencia 2009. "Islands of Effective International Adjudication: Constructing an Intellectual Property Rule of Law in the Andean Community," *The American Journal of International Law* 103(1): 1-47.
- Helfer, Laurence R. and Voeten, Erik 2014. "International Courts as Agents of Legal Change: Evidence from LGBT Rights in Europe," *International Organization* 68(1): 77-110.
- Helfer, Laurence R. and Slaughter, Anne-Marie 1997. "Towards a Theory of Effective Supranational Adjudication," *Yale Law Journal* 107(2): 273-391.
- 2005. "Why States Create International Tribunals: A Response to Professors Posner and Yoo," *California Law Review* 93(3): 899-956.
- Hovi, Jon, Sprinz, Detlef and Underdal, Arild 2003. "The Oslo-Potsdam Solution to Measuring Regime Effectiveness," *Global Environmental Politics* 3(3): 74-96.
- Keohane, Robert O., Moravcsik, Andrew and Slaughter, Anne-Marie 2000. "Legalized Dispute Resolution: Interstate and Transnational," *International Organization* 54(3): 457-488.
- Kim, Hunjoon and Sikkink, Kathryn 2010. "Explaining the Deterrence Effect of Human Rights Prosecutions for Transitional Countries," *International Studies Quarterly* 54(4): 939-963.

- Kim, Moonhawk 2008. "Costly Procedures: Divergent Effects of Legalization in the GATT/WTO Dispute Settlement Procedure," *International Studies Quarterly* 52(3): 657-686.
- Lutz, Ellen and Sikkink, Kathryn 2001. "The Justice Cascade: The Evolution and Impact of Foreign Human Rights Trials in Latin America," *Chicago Journal of International Law* 2(1): 1-33.
- Mattli, Walter and Slaughter, Anne-Marie 1998. "Revisiting the European Court of Justice," *International Organization* 52(1): 177-209.
- Miles, Edward L., Underdal, Arild, Andresen, Steinar, Wettestad, Jørgen, Skjærseth, Jon Birger and Carlin, Elaine M. (eds.) 2002. *Environmental Regime Effectiveness: Confronting Theory with Evidence*. Cambridge, MA: MIT Press.
- Mitchell, Ronald 2008. "Evaluating the Performance of Environmental Institutions: What to Evaluate and How to Evaluate It?," in Young, Oran R., King, Leslie and Schroeder, Heike (eds.), *Institutions and Environmental Change: Principal Findings, Applications and Research Frontiers*. Cambridge, MA: MIT Press, pp. 79-114.
- Posner, Eric and Yoo, John 2005. "Judicial Independence in International Tribunals," *California Law Review* 93(1): 1-74.
- Rodman, Kenneth A. 2008. "Darfur and the Limits of Legal Deterrence," *Human Rights Quarterly* 30(3): 529-560.
- Romano, Cesare P.R. 2011. "A Taxonomy of International Rule of Law Institutions," *Journal of International Dispute Settlement* 2(1): 241-277.
- 2014. "Trial and Error in International Judicialization," in Romano, Cesare P.R., Alter, Karen J. and Shany, Yuval (eds.), *The Oxford Handbook of International Adjudication*. Oxford: Oxford University Press, pp. 111-134.
- Romano, Cesare P.R., Alter, Karen J. and Shany, Yuval 2014. "Mapping Interantional Adjudicative Bodies, the Issues, and Players," in Romano, Cesare P.R., Alter, Karen J. and Shany, Yuval (eds.), *The Oxford Handbook of International Adjudication*. Oxford: Oxford University Press, pp. 3-26.
- Rosendorff, R. Peter 2005. "Stability and Rigidity: Politics and Design of the WTO's Dispute Settlement Procedure," *American Political Science Review* 99(3): 389-400.
- Shany, Yuval 2014. *Assessing the Effectiveness of International Courts*. Oxford: Oxford University Press.
- Shelton, Dinah L. 2009. "Form, Function, and the Power of International Courts," *Chicago Journal of International Law* 9(2): 537-571.
- Sikkink, Kathryn 2011. *The Justice Cascade: How Human Rights Prosecutions Are Changing World Politics*. New York: W.W. Norton & Company.
- Simmons, Beth A. 2009. *Mobilizing for Human Rights: International Law in Domestic Politics*. Cambridge: Cambridge University Press.
- Stone Sweet, Alec 2004. *The Judicial Construction of Europe*. Oxford: Oxford University Press.
- Stone Sweet, Alec and Brunell, Thomas L. 1998. "Constructing a Supranational Constitution: Dispute Resolution and Governance in the European Community," *The American Political Science Review* 92(1): 63-81.
- 2012. "The European Court of Justice, State Noncompliance, and the Politics of Override," *American Political Science Review* 106(1): 204-213.
- 2013. "Trustee Courts and the Judicialization of International Regimes. The Politics of Majoritarian Activism in the European Convention on Human Rights, the European Union, and the World Trade Organization," *Journal of Law and Courts* 1(1): 61-88.

- Tallberg, Jonas, Sommerer, Thomas, Squatrito, Theresa and Lundgren, Magnus 2016. "The Performance of International Organizations: A Policy Output Approach," *Journal of European Public Policy* 23(7): 1077-1096.
- Wippman, David 1999. "Atrocities, Deterrence, and the Limits of International Justice," *Fordham International Law Journal* 23(2): 473-488.
- Young, Oran R. (ed.) 1999. *The Effectiveness of International Environmental Regimes: Causal Connections and Behavioral Mechanisms*, Cambridge, MA: MIT Press.
- Zangl, Bernhard 2008. "Judicialization Matters! A Comparison of Dispute Settlement under Gatt and the WTO," *International Studies Quarterly* 52(4): 825-854.
- Zangl, Bernhard, Helmedach, Achim, Mondré, Aletta, Kocks, Alexander, Neubauer, Gerald and Blome, Kerstin 2012. "Between Law and Politics: Explaining International Dispute Settlement Behavior," *European Journal of International Relations* 18(2): 369-401.