

University of Chicago Law School

Chicago Unbound

Journal Articles

Faculty Scholarship

2005

Reply to Helfer and Slaughter

Eric A. Posner

John C. Yoo

Follow this and additional works at: https://chicagounbound.uchicago.edu/journal_articles



Part of the [Law Commons](#)

Recommended Citation

Eric Posner & John C. Yoo, "Reply to Helfer and Slaughter," 93 California Law Review 957 (2005).

This Article is brought to you for free and open access by the Faculty Scholarship at Chicago Unbound. It has been accepted for inclusion in Journal Articles by an authorized administrator of Chicago Unbound. For more information, please contact unbound@law.uchicago.edu.

Reply to Helfer and Slaughter

Eric A. Posner[†] and John C. Yoo[‡]

In *Judicial Independence in International Tribunals*, we evaluated the argument that successful international dispute resolution depends on the independence of international judges.¹ On this view, independent international courts are more likely to deliver impartial judgments, to be used by states, and to enjoy a high rate of compliance with their judgments. Our paper first developed a theory that illustrated the tradeoffs posed by independence: states gain a more consistent jurisprudence and avoid the costs of constructing ad hoc tribunals for specific disputes, but they lose control over judges who might decide cases in a manner that is inconsistent with the states' (unilateral or joint) interests. We then attempted to test the hypothesis that independence and effectiveness are positively correlated. We concluded that the evidence for such a correlation is weak and that the correlation may in fact be negative.

Our Article acknowledged that the European courts appear to be effective. But we argued that Europe presents special circumstances because of its half-century of political integration. Because the international system lacks Europe's level of political integration, judges on international courts have less room to maneuver than judges on European courts, and if they take advantage of agency slack to pursue agendas that conflict with state interests, their courts are more likely to be ineffective. We concluded that the new generation of international courts, including the International Criminal Court (ICC) and the International Tribunal for the Law of the Sea (ITLOS), may fail to accomplish their original goals because their judges have too much independence.

Because of our paper's empirical orientation, it did not discuss Helfer and Slaughter's earlier work, which made normative arguments. We did cite it, however, because we thought its empirical assumptions, which reflected the general view of international law scholars, were inconsistent

Copyright © 2005 California Law Review, Inc. California Law Review, Inc. (CLR) is a California nonprofit corporation. CLR and the authors are solely responsible for the content of their publications.

[†] Kirkland & Ellis Professor of Law, University of Chicago.

[‡] Professor of Law, School of Law, University of California, Berkeley (Boalt Hall); Visiting Scholar, American Enterprise Institute. We thank Wayne Hsiung and Miguel de Figueiredo for very helpful comments. Posner thanks the Russell Baker Scholars Fund, and Yoo thanks the Boalt Hall Fund for financial support.

1. Eric A. Posner & John C. Yoo, *Judicial Independence in International Tribunals*, 93 CALIF. L. REV. 1 (2005).

with our findings. In their response to *Judicial Independence*, Helfer and Slaughter have devoted much of their time to amplifying their earlier theories. In this reply, we do not address that effort but only respond to specific arguments that they have made about our piece.

I THEORY

Our earlier Article sought to understand the design features of successful and unsuccessful international tribunals.² We based our thesis on a straightforward principal-agent approach. States (principals) establish tribunals (agents) in order to serve their interests—here, to resolve disputes that arise because of disagreements about facts and law. The problem for states is that if tribunals are to function as intended, states must agree in advance to abide by the tribunals' decisions. By doing so, however, they risk committing themselves to decisions that are, as we put it, outside the states' "win set." Tribunals need a minimum level of independence in order to do their job properly. The traditional arbitration system provides for such minimal independence with its requirement that the arbitrators appointed by each state jointly agree on a third, tie-breaking arbitrator.

Adjudicators comply with the interests of the states that appoint them for various reasons. In the traditional arbitration system, the tie-breaking arbitrator would not be reappointed in subsequent arbitrations unless the party arbitrators believed that he or she resolved the dispute consistent with the interests of the disputing states. By contrast, in modern international tribunals, a judge may not be reappointed at the end of his or her term. This latter sanction is weaker because the terms of judges on modern international tribunals are, by design, relatively long and because the votes of all the judges may dilute the vote of a single judge. If, for example, all the judges on the International Court of Justice (ICJ) rule against state X, state X cannot sanction them because it does not have control over their reappointments; at best, state X might have influence over the reappointment of only one judge, though more likely over none. Thus, state X's only remedies are to withdraw from the tribunal or the treaty regime, which may be costly, or to seek changes to the regime, which will require the consent of other states.

In our Article we identified several benefits that would be created by independent tribunals if their impartiality could be guaranteed.³ The most obvious benefit is that, by contrast to the traditional arbitration system, which can resolve only bilateral disputes, an independent tribunal could resolve disputes among multiple parties or disputes in which multiple parties have an interest in the outcome. In a multilateral treaty regime, third

2. *Id.*

3. *Id.* at 20-21.

parties to a dispute might have an interest in the development of the law by a tribunal, and neutral development of the law might be possible only through a multilateral tribunal with substantial judicial independence.

Thus, a tradeoff exists, and it is certainly possible as a matter of theory that the benefits of judicial independence exceed the agency costs. In our view, however, the international law literature focuses too much on the advantages of judicial independence and neglects the concerns about agency costs. This strikes us as a mistake. Consider the following analogy. While the world's problems could, in principle, be solved through a world government, states have not yet created such an institution, and virtually no one thinks such an institution is possible in the imaginable future. Governments fear that any world government, however organized, would not be responsive to their concerns and those of their citizens. In the jargon, states do not grant powers to a world government because of agency costs. But courts are just one feature of a government. The puzzle for supporters of independent international tribunals is why an international judicial branch might be possible if international executive and legislative branches are not.

Our conjecture was that these three branches are not so different, and what applies to the latter two applies to the former as well. If world executive and legislative institutions have been weak, then world courts should be as well. This led us to ask whether international tribunals have actually been ineffective. To study the effectiveness of international adjudication we took the one variable that seems most characteristic of domestic judicial decision making—independence—and analyzed the extent to which it correlates with judicial effectiveness, which we defined as frequent usage of a tribunal and high compliance with its decisions. We identified a spectrum of attributes that made international courts dependent or independent: the method of choosing judges, the nature of their jurisdictions, the identity of the parties to a dispute, the number of state parties, and the lifespan of the tribunal. We then examined evidence that would show whether greater independence increases usage, compliance, or effectiveness.

Helfer and Slaughter do not identify any defects in our principal-agent framework. Instead, they emphasize the role of international courts as a device that helps states cooperate by enhancing the strength of their commitments.⁴ We do not disagree with this premise, which has been advanced previously in the international relations literature; indeed, it is assumed by our theory. More precisely, however, the mechanism that allows states to make commitments to adhere to their promises is their interest in preserving reputation, or their fear of retaliation. It is not adjudication.

4. Laurence R. Helfer & Anne-Marie Slaughter, *Why States Create International Tribunals: A Response to Professors Posner and Yoo*, 93 CALIF. L. REV. 899, 902 (2005) [hereinafter, Helfer & Slaughter, *Response*].

Adjudication works only if states are willing to comply with the legal obligation that sets up the adjudicatory mechanism in the first place and commits the parties to comply with adverse judgments. Adjudication itself only adds information. Even if states are more willing to commit themselves to a treaty regime with an adjudicatory mechanism than a regime that lacks one, the adjudicatory mechanism does not, by itself, provide a reason for a state to comply with a commitment. If states did not already care about their reputation for complying with a treaty's obligations, then they could not, by setting up international courts, commit themselves to anything. Our disagreement is not about whether international tribunals are precommitment devices, but whether they work effectively when judges are independent. After all, if only dependent tribunals are effective, then the value of independent international tribunals as precommitment devices is limited.

Helfer and Slaughter attempt to distinguish their approach from our principal-agent framework by relying on the idea that international judges are more like "trustees" than "agents." It is unclear why trustees are to be distinguished from agents; trustees are simply a subspecies of agent with a certain amount of discretion and independence. Trustees still work for a principal, just as agents do. Helfer and Slaughter borrow this distinction from political scientist Karen Alter, who argues that international judges may pursue an agenda that is independent of the immediate, political interests of the states that have appointed them.⁵ We interpret Alter's work as saying that states expect international judges to be impartial, and that international judges take this expectation seriously and try to be impartial. Alter understands, however, that states may be as suspicious of "trustees" as of "agents," and therefore may be reluctant to delegate power to independent tribunals.⁶ So we are back where we started, with the question of whether courts, as "agents" or "trustees," are actually impartial. We see no reason for Alter's terminological innovation, and in any event do not understand why Helfer and Slaughter think that it supports their argument.

To the extent that Helfer and Slaughter develop a different theory, they argue that we mistakenly focused on states. They argue that the correct unit of analysis is the subnational actor. Although they attempt to explain how such actors might exert an impact on the creation of international tribunals, they do not develop a theory, assemble any data, or analyze our

5. See Karen J. Alter, *Delegation to International Courts: Four Varieties and Their Implications for State-Court Relations* (2003) (unpublished manuscript) (on file with authors) [hereinafter, Alter, *Delegation*]; see also Karen J. Alter, *Agents or Trustees? International Courts in their Political Context* (2004) (unpublished manuscript) (on file with authors) [hereinafter, Alter, *Agents or Trustees?*]; Helfer & Slaughter, *Response*, *supra* note 4, at 932 n.110 (citing Alter, *Delegation*).

6. See Alter, *Agents or Trustees?*, *supra* note 5, at 6-7. Elsewhere, Alter argues that WTO panelists should be appointed to the panels on the basis of their political sensitivity. See Karen J. Alter, *Resolving or Exacerbating Disputes? The WTO's New Dispute Resolution System*, 79 INT'L AFF. 783, 798 (2003).

data with this link in mind. One might hypothesize, for example, that greater participation of interest groups leads to greater state usage of certain international tribunals and more compliance with their decisions. One could then attempt to measure interest group activity and test whether it correlates with tribunal effectiveness. Neither Helfer and Slaughter's original piece, nor their Response to our article, undertakes such a project.

What, then, was Helfer and Slaughter's original paper about? It advanced a normative argument that the United Nations Human Rights Committee should be transformed from a committee—whose only power is to issue reports about the human rights practices of various countries—into a court.⁷ The authors claimed to “have developed the . . . compilation of the attributes of effective supranational adjudication by distilling commentary and analysis by judges, lawyers, and political scientists who have closely observed the workings of the ECJ and the ECHR, supplementing these findings with [their] own analysis.”⁸ They then presented a narrative description of the attributes of the ECJ and the ECHR, summarizing their characteristics in a “checklist” of thirteen factors. Other international courts that also bear these attributes, Helfer and Slaughter argued, would also be effective.⁹ They concluded that the UN Human Rights Committee would be more effective if it were transformed into a tribunal characterized by their thirteen factors.¹⁰

Helfer and Slaughter premised their argument on the theory that the ECJ and the ECHR are effective, and then reasoned that international institutions that mimic their characteristics will be similarly effective. This was a mistake. The ECJ and ECHR are not representative of the world's international tribunals. A larger sample of tribunals would have provided a firmer basis for evaluating the proposal that the UN Human Rights Committee be converted into a tribunal.

II

INTERPRETING THE DATA

Helfer and Slaughter accuse us of committing several methodological errors. We classify them as follows: (a) omitted variable bias, (b) selection bias, and (c) measurement error.

7. Laurence R. Helfer & Anne-Marie Slaughter, *Toward a Theory of Effective Supranational Adjudication*, 107 *YALE L.J.* 273, 298 (1997) [hereinafter Helfer & Slaughter, *Effective Supranational Adjudication*].

8. *Id.* at 298.

9. *Id.* at 336-37.

10. *Id.* at 337-38.

A. Omitted Variable Bias

Helfer and Slaughter claim that the list of variables that we omitted in the development of our theory is long.¹¹ They argue that we should have considered variables that are more or less the ones that make up their earlier “checklist” for creating effective independent tribunals.¹² For example, they claim that we did not control for judicial access or “political and discursive constraints that . . . affect a tribunal’s perception of its authority in relation to the states that established it.”¹³

Omitted variable bias occurs when an apparent correlation between a dependent variable and an independent variable actually reflects the correlation between the dependent variable and another independent variable that happens to be (partially) correlated with the independent variable that was used.¹⁴ To illustrate, a study might find a correlation between drinking wine and lower rates of heart disease. Such a study may not take into account that wealth is also correlated with lower heart disease and with drinking more wine. To fix this problem, the authors of the study would add wealth as a control variable. The new study might then show that wealthy people are healthy and drink more wine, not that drinking wine improves one’s health.

To show that our strong thesis suffers from omitted variable bias,¹⁵ Helfer and Slaughter would have to claim that omitted variable bias undermines the claim that independence causes ineffectiveness. For example, if our failure to include access as a variable were an example of omitted variable bias, then inclusion of the variable would have shown that access (rather than independence) was correlated with ineffectiveness. But Helfer and Slaughter do not think that access causes ineffectiveness; they think that access causes, or contributes to, effectiveness.¹⁶ They have omitted variable bias backwards.

To demonstrate omitted variable bias, Helfer and Slaughter must show an independent variable that causes *ineffectiveness* and which is also correlated with independence. For example, omitted variable bias might exist if it could be shown that militarily powerful countries tend to create independent tribunals but use them and comply with their judgments less often than do other countries. Helfer and Slaughter do not make any such argument and so provide no reason to think that omitted variable bias is a problem for our study.

11. Helfer & Slaughter, *Response*, *supra* note 4, at 928.

12. *Id.*

13. *Id.*

14. GARY KING, ROBERT KEOHANE, & SIDNEY VERBA, *DESIGNING SOCIAL INQUIRY* 169 (1994).

15. There would be no point in arguing that the weak thesis suffers from omitted variable bias, as it argues that there is no correlation in the first place.

16. Helfer & Slaughter, *Response*, *supra* note 4, at 928-29.

Helfer and Slaughter do not really think that we committed omitted variable bias; they are confusing this problem with other methodological problems that we will discuss shortly. But their argument does raise a larger issue, which is why we did not use more explanatory variables. The purpose of conducting an empirical study is to identify relevant independent variables that have a causal relationship with a dependent variable. One problem for such studies is isolating and defining the right independent variables in advance; one can always argue that there is some other independent variable that has an impact on the dependent variable. But in an effort to find the most parsimonious explanation possible, we (like all researchers) must draw a line somewhere. This decision is based on a judgment of what variables are sufficiently relevant to justify study, and which are not important enough and should be excluded. To take our wine-heart disease example again, one could always criticize such a study by claiming that it did not include all of the possible causes of heart disease, some of them more tenuous than others, such as living conditions, geographic location, commuting times, number of pets, and so on. The researcher must make a judgment about which variables are important to include and which are not, because otherwise a study loses its power to differentiate among alternative hypotheses.¹⁷

The right way to choose independent variables is to start with a theory. In our case, we started with a principal-agent model that drew attention to the problem of controlling agents. This model immediately raises the question of whether judicial independence is desirable; hence, our use of it as the explanatory variable. The wrong way to choose independent variables can be illustrated by Helfer and Slaughter's earlier paper. That paper did not provide a theory explaining why the thirteen attributes that they identified were responsible for the ECJ's and ECHR's effectiveness—as opposed to any of the large set of other attributes that these two courts share.

B. Selection Bias

Helfer and Slaughter claim that we have committed selection bias by focusing on eleven international courts, rather than including the additional fifteen that they list.¹⁸ We did not include these additional fifteen tribunals for several straightforward reasons.

First, three of the international tribunals whose omission Helfer and Slaughter cite are not even in operation.¹⁹ There are simply no data on the dependent variable.

17. For a discussion of this problem, see King et al., *supra* note 14, at 118-21.

18. *Id.*

19. The Court of Justice of the African Union (CJAU), the Caribbean Court of Justice (CCJ), and the African Court of Human and Peoples' Rights (ACHPR).

Second, two of the courts are international criminal tribunals established by command of the United Nations Security Council.²⁰ Their jurisdictions apply to individuals, not to states. They do not resolve disputes between nations, or even between nations and individuals, and thus fall outside the parameters of our earlier article. In any event, we doubt that their inclusion would have supported Helfer and Slaughter's argument rather than our conclusions. Many have criticized the international criminal tribunals for being slow, expensive, and ineffective.²¹ But our main point is that data from these tribunals are not relevant to the question of whether *states* will more often submit to independent or dependent tribunals and comply with their decisions.

Third, data on the remaining ten courts on Helfer and Slaughter's list are, as far as we can tell, poor or inaccessible. We omitted these tribunals from our study because the basic principle that investigators should not use irrelevant or unreliable data is just as important as avoiding selection bias. We conducted a search for reported decisions of these tribunals when we originally gathered our data and decided that we could not rely on the limited information available. Indeed, the work by Karen Alter, from which Helfer and Slaughter have drawn upon so much for data and theory, acknowledges this very point.²² We chose to be cautious and to assume that bad data could have an innocent explanation. It is possible that an effective tribunal does not keep public records, or that scholars have not evaluated its performance.

Another possibility, however, is that a tribunal generates unreliable data because it is ineffective. This possibility is illustrated by the Central American Court of Justice (CACJ). The CACJ was created by the 1991 Protocol of Tegucigalpa to the Charter of the Organization of Central American States, which itself established an institutional framework for the

20. The International Criminal Tribunal for Rwanda (ICTR) and the International Criminal Tribunal for the Former Yugoslavia (ICTY).

21. See, e.g., James Blount Griffin, *Note: A Predictive Framework for the Effectiveness of International Criminal Tribunals*, 34 VAND. J. TRANSNAT'L L. 405, 432 (2001) (calling the ICTY "too slow for the demands of justice").

22. Karen Alter's paper on international tribunals lists the same courts as Helfer and Slaughter. Alter, *Delegation*, *supra* note 5, at 5. She reports that the Benelux Court, the Economic Court of the Commonwealth of Independent States, the African Court of Human and Peoples Rights, the Court of Justice for the Common Market of Eastern and Southern Africa, the Court of Justice for the Arab Magreb Union, and the International Criminal Tribunal for Sierra Leone have either no cases or no data available. The Judicial Tribunal for the Organization of Arab Petroleum Exporting Countries (two cases) and the Common Court of Justice and Arbitration for the Organization for the Harmonization of Corporate Law in Africa (four cases) have far too little data to be relevant. In addition, we have found that the Court of Justice for the Andean Community has little reliable data available. See, e.g., COURT OF JUSTICE OF THE ANDEAN COMMUNITY, GENERAL INFORMATION, at <http://www.comunidadandina.org/ingles/who/court.htm> (last visited Jan. 27, 2005). We have found no rigorous academic studies of any of these courts.

economic and political integration of Central America.²³ Helfer and Slaughter list the CACJ as an independent tribunal whose exclusion creates selection bias,²⁴ but they do not provide any data concerning the dependent variable—the compliance and usage levels for the CACJ.²⁵ We did not include the CACJ in our original paper because of a lack of reliable data on the dependent variable.²⁶

Closer examination shows that the reason for the lack of reliable data may be that the court has been unsuccessful. Even though the Presidents of the five Central American countries—Costa Rica, El Salvador, Guatemala, Honduras, Nicaragua—signed the Protocol of Tegucigalpa and the Statute of the CACJ, only three countries—El Salvador, Honduras, and Nicaragua—have ratified it.²⁷ From the few official reports, it appears that the number of cases is in the single digits; one commentator describes the docket as “light.”²⁸ In terms of usage, therefore, it would be easy to conclude that the CACJ has been unsuccessful.

Anecdotal evidence regarding compliance reinforces this conclusion. For example, in 1999, Nicaragua sued Honduras in the CACJ for ratifying a treaty with Colombia (a nonparty to the Protocol or the CACJ) over waters and islands claimed by Nicaragua.²⁹ The CACJ issued a preliminary order demanding that Honduras suspend its ratification of the agreement. Nicaragua also responded to Honduras’s actions by imposing a 35% duty on Honduran goods. Honduras responded with its own suit in the CACJ claiming that Nicaragua’s retaliation had violated its treaty obligations. The CACJ issued another preliminary order that Nicaragua suspend the duty. Neither Nicaragua nor Honduras obeyed the CACJ’s orders.³⁰ To make matters worse, Honduras suspended its participation in the CACJ in August 2004.³¹ These are not the signs of an effective court that is

23. PROTOCOL OF TEGUCIGALPA TO THE ORGANIZATIONAL LETTER OF THE ORGANIZATION OF CENTRAL AMERICAN STATES (Dec. 13, 1991), at <http://www.sice.oas.org/trade/sica/SG121391.asp>.

24. Helfer & Slaughter, *Response*, *supra* note 4, at 912 tbl. 1. (giving the tribunal four of the five possible points in the independence scale we outlined in our Article).

25. We are unsure whether Helfer and Slaughter have correctly scored the independence of the CACJ. While they state that the court does not require that each nation have a member on the court, it appears that this is in fact the rule. See STATUTE OF THE CENTRAL AMERICAN COURT OF JUSTICE, arts. 9-11, 34 I.L.M. 923 (1992); see also Thomas A. O’Keefe, *The Central American Integration System at the Dawn of a New Century*, 13 FLA. J. INT’L L. 243 (2001).

26. The Court’s reports, available in Spanish on its website, show an extremely low number of cases (we do not include advisory opinion requests, which were also low), at <http://www.ccj.org.ni> (last visited Jan. 27, 2005). By our count, they appear to refer to perhaps fifteen contentious cases; several resolutions involve only one case, and some resolutions involve advisory opinions. <http://www.ccj.org.ni> (last visited Jan. 27, 2005). Again, we think that these numbers are too small to be worth including in our study.

27. See O’Keefe, *supra* note 25, at 251.

28. *Id.* at 253.

29. *Id.* at 243.

30. *Id.* at 254-55.

31. The letter can be found at <http://www.ccj.org.ni>.

experiencing compliance with its decisions. We could have expanded our empirical study to such courts as the CACJ. Their low levels of usage and the lack of compliance with their decisions might actually have supported our thesis. But to have included these anecdotes and sought to determine the court's performance without reliable reports of its decisions would risk the introduction of unreliable data.

Helfer and Slaughter also cite the European Free Trade Area Court (EFTAC)—which hears cases between Norway, Iceland, and Liechtenstein—and the Benelux Court of Justice (BCJ)—which hears cases between Belgium, Luxembourg, and the Netherlands—as courts that we should have included in our article. Inclusion of data on these courts would not have made sense, however. If Helfer and Slaughter were right that we incorrectly excluded the ECJ, then the ECJ itself should stand as a sufficient refutation of our hypothesis. If we are right, then our reasons for distinguishing the ECJ and the ECHR would apply to all European-based tribunals, such as the EFTAC and the BCJ.

This brings us to our differences with Helfer and Slaughter concerning the significance of the ECJ and the ECHR. *Judicial Independence* acknowledged that the ECJ appears to be effective based on our measures of usage and compliance. The ECJ, however, is neither a national nor an international tribunal. Because Europe is a quasi-confederation, its tribunals cannot be easily classified.³² Hence, the ECJ cannot serve as a clean piece of evidence for the hypothesis that judicial independence and effectiveness are correlated.

Helfer and Slaughter claim that we excluded Europe to fit our results.³³ They suggest that we should not have inserted a new variable—political unification, in this case—when the data set is not large enough to subject the new variable to a test of significance. But the problem is as much theirs as ours. Recall their argument that the UN Committee on Human Rights should be converted into an international court. They originally developed their checklist to show how such nonjudicial, essentially advisory bodies, could be transformed. To work, however, a UN Human Rights “Court” would require jurisdiction over (nearly) the entire world, not just the advanced European countries. There is no reason to think that a court that works for Europe, where political and legal institutions in most countries are of high quality, would work for a world political community that lacks the same level of cohesion and integration. Whatever one thinks about the EU, it is nothing like the international community.

One possible conclusion to draw from Helfer and Slaughter's criticisms is that there is just not enough data to test the hypothesis that judicial

32. Cf. J.H.H. Weiler, *The Transformation of Europe*, 100 YALE L.J. 2403, 2413 (1991) (observing that Europe was becoming a “federal state”).

33. Helfer & Slaughter, *Response*, *supra* note 4, at 917-22.

independence and effectiveness are correlated. That is, in fact, our weak thesis and may be the lesson of this debate for those who are uneasy about drawing conclusions based on data sets without a large number of observations. But if this is a problem for us, it is far worse for Helfer and Slaughter. In their earlier piece, they identified thirteen variables from an examination of only two European courts and did not provide any statistics or empirical evidence to support their claim of a correlation between independence and effectiveness. Helfer and Slaughter's response now demands that numerous other factors—interest groups, domestic political institutions, whether the treaty is bilateral or multilateral, “deep” or “shallow,” and so forth—be included as independent variables. But if the low number of observations is the problem, multiplying the number of independent variables only aggravates it.

Helfer and Slaughter also argue that we have committed selection bias by excluding what they call “quasi-judicial bodies” that review compliance with international law but are not considered international courts.³⁴ However, there is a significant difference between bodies such as the ICJ, the ITLOS, or the GATT/WTO on the one hand, and entities such as the UN Human Rights Committee or even the Inter-American Human Rights Commission on the other. Treaties vest international courts with the authority to issue decisions that are binding as a matter of international law.³⁵ They resolve disputes through cases brought by states in a dispute and base their decisions on the interpretation of a treaty or other source of international law.

Entities such as the UNHRC, however, can only issue reports or make recommendations that nations may ignore without violating international law.³⁶ They do not hear cases between parties, nor do they issue binding legal judgments, and they have no “jurisdiction” in the conventional sense. In *Judicial Independence* we had to set some limit to our study and decide what institutions were worth studying. Otherwise, why stop at quasi-judicial bodies? Why not also include nongovernmental organizations, like Amnesty International, or national entities, like the U.S. State Department, both of which also make comments and issue reports on state compliance with international law?

34. *Id.* at 923.

35. In the case of GATT, decisions are binding only if the panel judgment is adopted by the members. This makes GATT a middle case, but like Slaughter herself we think it sensible to classify the GATT panel system as an adjudicatory system. See Robert O. Keohane, Andrew Moracsik, & Anne-Marie Slaughter, *Legalized Dispute Resolution: Interstate and Transnational*, 54 INT'L ORG. 457, 461 (2000).

36. See generally, OFFICE OF THE HIGH COMMISSIONER FOR HUMAN RIGHTS, INTRODUCTION TO THE HUMAN RIGHTS COMMITTEE, available at <http://www.unhchr.ch/html/menu2/6/a/introhrc.htm> (last visited Jan. 27, 2005) (setting out a basic structure of the UNHRC).

C. Measurement Error

Helfer and Slaughter argue that we have improperly measured our two main variables, independence and effectiveness.³⁷ They make no effort to develop alternate, workable means of measuring these variables, nor do they provide any new data of their own. We believe that our approach still represents the best way so far of measuring these variables.

Helfer and Slaughter criticize us for measuring only independence and ignoring the other factors on their checklist.³⁸ Most of their factors, however, are too poorly defined to measure. Just to take three of many possible examples from their checklist, Helfer and Slaughter believe that “awareness of audience,” “incrementalism,” and “judicial cross-fertilization and dialogue” are variables.³⁹ They do not make a serious effort to define what these terms mean, and they never attempt to measure them. In their reply, they expand the variables to include “political and discursive constraints.”⁴⁰ Again, they do not explain how to measure this variable. They also argue that we should have looked at substance rather than form by measuring whether judges are “really” independent rather than only formally so.⁴¹ The pitfalls of this suggestion are evident in their own discussion of independence. For example, Helfer and Slaughter state that compliance with a judgment is a measure of independence. This explanation collapses the independent and dependent variables, making empirical verification of any causal relationship impossible.

Our study is certainly open to criticism for not fully capturing judicial independence. But constructive criticism would provide an alternative approach. Helfer and Slaughter fail to explain how one can measure their proposed variables in a nonarbitrary way. Instead, they simply assert that if we had used their thirteen-point index, we would have found that more independent tribunals are more effective. We doubt this, and we challenge them to conduct such an empirical study. We believe that they will find it impossible to define and measure many of the thirteen variables in their checklist. In any event, an empirical study that uses thirteen explanatory variables on a data set consisting of a dozen or so observations cannot yield determinate results.

Helfer and Slaughter discuss the difficulties with measuring effectiveness, which we described and acknowledged in our paper. There is no easy response to the challenge of measuring this complicated phenomenon, and we went to great pains in our earlier article to discuss the advantages and

37. Helfer & Slaughter, *Response*, *supra* note 4, at 909-31.

38. *Id.* at 902.

39. Helfer & Slaughter, *Effective Supranational Adjudication*, *supra* note 7, at 308-24.

40. Helfer & Slaughter, *Response*, *supra* note 4, at 930.

41. *Id.* at 931.

disadvantages of the various approaches.⁴² We do not believe that they have introduced any alternative method for measuring usage or compliance.

D. *The Burden of Proof*

Finally, we note a pervasive error in Helfer and Slaughter's Response. They claim that even our weak thesis is unpersuasive because of the various problems of selection bias, omitted variable bias, and measurement error,⁴³ but misunderstand the implications of this claim. If it is true that selection bias, measurement error, and the other problems are decisive, then we are, in fact, left with our weak thesis: there is no evidence of any relationship between tribunal independence and effectiveness (positive or negative). A fortiori, there can be no evidence of any relationship between their thirteen factors and tribunal effectiveness. To provide evidence for a claim of a positive relationship between their various factors and tribunal effectiveness, Helfer and Slaughter must *overcome* the statistical problems they identify.

Why is it important to advance the weak thesis? First, because among international law scholars the impartiality and effectiveness of independent international tribunals are articles of faith. If we are right, their elaborate and ambitious normative arguments are constructed on foundations of sand. Second, because states continue to create independent tribunals—the topic of the next section.

V

WHY DO STATES CONTINUE TO CREATE INDEPENDENT TRIBUNALS?

Helfer and Slaughter's strongest claim is that states would not continue to create independent tribunals, such as the ICC and the ITLOS, if either our weak thesis or our strong thesis were correct. We speculated that states may have thought the gains from independence exceeded the costs but had been mistaken.⁴⁴ This is not inconsistent with our central premise: even rational actors make mistakes. In this case, government actors thought the rewards were worth the risks. If the institutions fail, there is no great loss, and if they succeed, much good could come of it. Over time, government actors will update their beliefs about the design of international tribunals and either stop creating independent tribunals or stop participating in the independent tribunals that they created.

This period of retrenchment may already be upon us. Although states continue to join the WTO and to seek entry to the European Union, we do not know whether those states do so to obtain the benefits of the

42. Posner & Yoo, *supra* note 1, at 27-29.

43. Helfer & Slaughter, *Response*, *supra* note 4, at 909-31.

44. Posner & Yoo, *supra* note 1, at 25-26.

adjudicatory institutions themselves, or simply the substantive benefits of the treaty regime in question.⁴⁵ Many of the modern tribunals (like the ICTY and the ICTR) do not put at risk the states that created them. The same can be said for the ICC, whose members consist mainly of states that do not expect that their citizens will commit war crimes or human rights violations on foreign soil. States that foresee a need to engage in significant military action, like the United States, have refused to join the ICC. Other tribunals, like ITLOS, have not been used much, and states seem reluctant to give it compulsory jurisdiction in the manner envisioned by its founders.⁴⁶ And when states create tribunals, they sometimes underfund or neglect them.⁴⁷ The puzzle, then, is not only why states continue to create international tribunals, but why states continue to create tribunals that they do not use very much.

One possible answer is that with the end of the Cold War, states have found themselves free to experiment with new international institutions. Tribunals are attractive for the theoretical reasons we discussed in our first paper.⁴⁸ But beyond that choice, determining the optimal design of a tribunal is complex, and it is not surprising that states have opted for more ambitious designs in the hope that they will end up with something like the ECJ. Participants are discovering the hazards of independent tribunals even before these tribunals begin operations and have begun pulling back.

Another possible answer is that states think there are valuable symbolic reasons for setting up tribunals that look like independent courts, and that by doing so they increase their prestige. States have been making symbolic international commitments for hundreds of years. If states are capable of entering a treaty that outlaws war, as did many of the world's powers in the Kellogg-Briand Pact, they are likewise capable of creating international tribunals that do little or nothing. Understanding how this symbolism works, who the audience is, and whether paper agreements may eventually acquire teeth, is a complicated problem, but enough historical evidence exists for this type of behavior that it cannot be dismissed out of hand.⁴⁹

45. Indeed, most of the growth in absolute caseload appears to come from the European tribunals, which we have acknowledged are effective tribunals but are not international in nature. See Posner & Yoo, *supra* note 1, at 53.

46. *Id.* at 70-72.

47. Indeed, Helfer and Slaughter cite an example themselves: the ICTY. Helfer & Slaughter, *Response*, *supra* note 4, at 948.

48. Posner & Yoo, *supra* note 1, at 25.

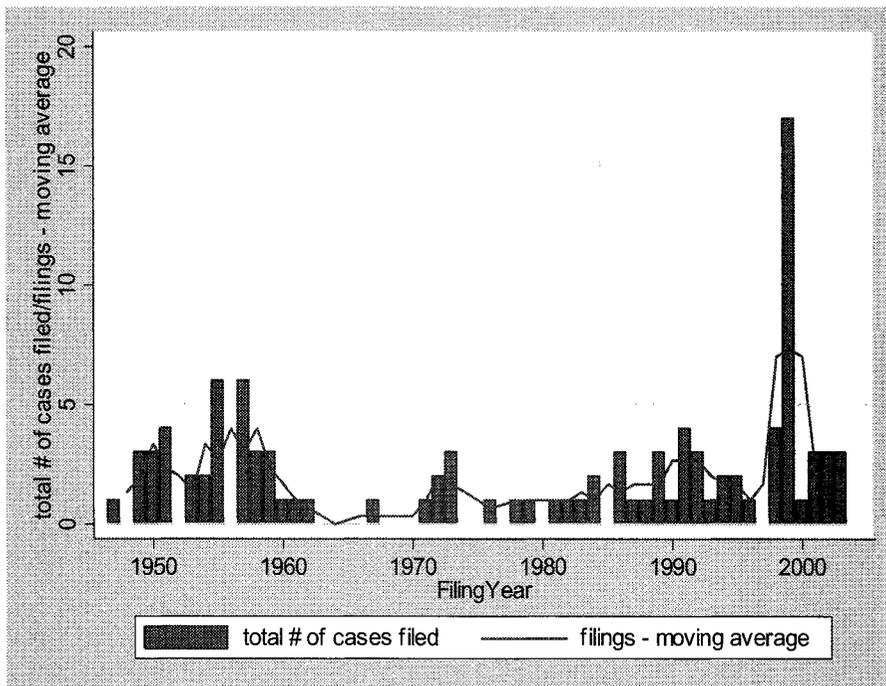
49. See Jack L. Goldsmith & Eric A. Posner, *Moral and Legal Rhetoric in International Relations: A Rational Choice Perspective*, 31 J. LEGAL STUD. 115 (2002) (discussing why states make symbolic commitments); this article is reprinted, in revised form, in JACK L. GOLDSMITH & ERIC A. POSNER, *THE LIMITS OF INTERNATIONAL LAW* 167-84 (2005).

CONCLUSION:
THE DECLINE OF THE ICJ

We close with a test case. While Helfer and Slaughter may disagree with us about the relevance of the ECJ and the ECHR, we all should be able to agree that the ICJ is relevant. It is the most prestigious and independent of the international courts, with the broadest jurisdiction and the greatest claim to legitimacy because of its age and the participation of nearly all states. Many people claim that the ICJ is enjoying a revival thanks to the end of the Cold War. The evidence suggests otherwise.⁵⁰

First, ICJ usage during the last twenty years is about one-third of what it was during its first twenty years, controlling for the increase in the number of states. In absolute terms, usage is about the same as it was during the ICJ's first twenty years (it was little used during the 1960s and 1970s). To the extent there was an increase, it was driven by Yugoslavia's 1999 complaint against ten NATO countries, which was treated as ten separate cases. Figure 1 provides the data.

Figure 1: ICJ Usage⁵¹



50. All of the information below is taken from Eric A. Posner, *The Decline of the International Court of Justice*, (2004) (unpublished manuscript) (on file with authors), available at http://papers.ssrn.com/sol3/papers.cfm?abstract_id=629341.

51. *Id.*

Second, major states (those with the ten largest economies) have all but stopped bringing complaints in the ICJ. In the last thirty years, they have filed only two complaints.⁵²

Third, the fraction of states that have consented to compulsory jurisdiction has declined from about two-thirds to one-third, and compulsory jurisdiction is now rarely the basis of litigation. Of the five permanent members of the UN Security Council, only the UK remains subject to compulsory jurisdiction.⁵³

Finally, the practice of conferring jurisdiction by treaty has also declined precipitously since the 1940s and 1950s. The number of treaties per year that contain clauses conferring jurisdiction on the ICJ has dropped from 9.7 during the ICJ's first twenty years,⁵⁴ to 1.3 during the ICJ's most recent twenty years. These are absolute numbers; remember that the number of states has tripled during this time period.

The one bright spot for the ICJ has been its special agreement cases, which confer jurisdiction *ex post*, in the manner of classical arbitration. Even here, however, states are increasingly insisting on four-judge chambers, which exclude most of the bench.⁵⁵

The puzzle for Helfer and Slaughter is that if independence contributes to effectiveness, why does all the evidence point to the decline of the ICJ? Why have states been avoiding it, restricting its jurisdiction, and, in special agreement cases, restricting the participation of its judges?⁵⁶ Our conjecture is that the ICJ has been a failure because the judges felt free to pursue agendas different from those of the states that set up and use the ICJ.⁵⁷ With no means for disciplining the judges, states have increasingly found it necessary to avoid the ICJ's jurisdiction and to limit its role in

52. *Id.*

53. *Id.*

54. The latter number might reflect date limitations, as it is unclear whether the data source includes the last few years. But the trend is clear enough: for the 1966 to 1985 period, there were only 2.8 treaties per year.

55. Posner, *supra* note 50, at 2.

56. See CHAYES & CHAYES, *THE NEW SOVEREIGNTY* 207 (1995) ("neither the [ICJ] nor other instruments of binding adjudicative settlement have shown themselves adaptable as instruments for the settlement of the stream of routine disputes that necessarily arise under a regulatory treaty regime"); see also W. MICHAEL REISMAN, *SYSTEMS OF CONTROL IN INTERNATIONAL ADJUDICATION AND ARBITRATION: BREAKDOWN AND REPAIR* 41-45 (1992) (describing the ICJ's problems).

57. Cf. REISMAN, *supra* note 56, at 41-43 (arguing that the judges feel beholden to the General Assembly when usage of the ICJ depends on the consent of litigants). For a valuable discussion of the view that the ICJ has a "progressive" agenda, see Michael Reisman, *Metamorphoses: Judge Shigeru Oda and the International Court of Justice*, 33 *CANADIAN YEARBOOK OF INT'L L.* 185 (1995). For evidence that ICJ judges vote the interests of their home states, see Eric A. Posner and Miguel de Figueiredo, *Is the International Court of Justice Biased?* (2004) (unpublished manuscript) (on file with authors), available at http://papers.ssrn.com/sol3/papers.cfm?abstract_id=642581.

their disputes. If so, the ICJ exemplifies the problems created when states grant independence to international judges.

