Rhetoric and Reality: International Organizations, Sovereignty Costs, and Human Rights

Emilie M. Hafner-Burton, Edward Mansfield and Jon Pevehouse
August 2019

Abstract

The number of international human rights institutions and countries participating in them has risen dramatically in recent decades, contributing to heated debates about their effectiveness in improving human rights. Much of that debate has focused on a small number of human rights treaties. Less attention has been paid to a larger set of international organizations (IOs) that claim to promote human rights. We argue that the extent to which these IOs improve human rights in member-states depends on their institutional design. IOs that intend to safeguard human rights must create costly commitments for members; otherwise, human rights rhetoric amounts to cheap talk. We assess our argument by evaluating the “sovereignty costs” associated with IOs marked by human rights mandates during the 1950 to 2012 period. We find that organizations that impose higher sovereignty costs tend to produce greater improvements in human rights than those that generate fewer sovereignty costs.

Word Count: 7,760
The number of international human rights institutions and countries participating in them has risen precipitously in recent decades, contributing to heated debates about their effectiveness.\(^1\) While scholars have devoted considerable effort to assessing the efficacy of a handful of human rights treaties, little attention has been paid to whether a larger and arguably more powerful set of international organizations (IOs) that also claim to promote, advance, or enforce human rights—what we will refer to as “human rights IOs”—have succeeded in enhancing the treatment of people in member-states. The purpose of this study is to address whether joining IOs with human rights mandates leads to improvement in member-states’ human rights practices and whether any such improvement is influenced by the extent to which organizations are designed to impose “sovereignty costs” on members.

We argue that IOs espousing human rights rhetoric must generate formal costly commitments on the part of participants in order to actually safeguard these rights. In the absence of such commitments, human rights rhetoric largely amounts to cheap talk. The costs of joining IOs depend on the sovereignty costs imposed on member-states; that is, the degree to which IOs require members to surrender policy autonomy regarding human rights in order to meet the institution’s rules, as well as how extensively the institution monitors those commitments. IOs characterized by high levels of legalism—which is promoted by rule specificity, issue linkage, membership restrictions, formal reporting, monitoring of behavior, and enforcement procedures—place more substantial constraints on members’ sovereignty.\(^2\) Joining these relatively demanding organizations should yield improvements in human rights on the part of new members, which often face both


international and domestic pressure to adjust their policies to conform to the organization’s standards. Other IOs, however, are designed to impose lower sovereignty costs on member-states—they create vague provisions, loopholes, low degrees of formal legal responsibility, and weak monitoring or oversight. Joining these IOs should lead to few changes in human rights behavior in new member governments.

Using a data set on human rights IOs and their associated sovereignty costs, we analyze how variation in institutional design influences member-states’ human rights behavior. Our results show that states joining IOs that impose higher sovereignty costs on members are more likely to improve human rights upon accession, whereas those entering IOs that impose fewer costs are less likely to improve such rights. Endemic to this line of research is the possibility that the states joining human rights IOs do not constitute a random sample. Our findings, however, are robust with respect to various strategies aimed at addressing any potential endogeneity in the analysis, including an instrumental variables approach. Further, there is evidence that improvements in human rights stemming from joining higher sovereignty cost IOs do not simply occur upon accession. Instead, they continue during each of the first five years of membership, before levelling off.

Our results indicate that, although existing debates over human rights institutions have focused on a small number of treaties and currently emphasize the domestic scope conditions necessary to stimulate compliance with those treaties, a wider range of international institutions that link rights to their other goals also influence governments’ treatment of their citizens. The ability of

---


4 Our data comes from Hafner-Burton, Mansfield, and Pevehouse 2015.

5 Downs, Rocke, and Barsoom 1996.
international institutions to safeguard human rights is therefore highly dependent on how they are designed.

**International Organizations and Human Rights**

Debates over the effectiveness of human rights institutions have grown heated in recent years. An initial wave of quantitative research found little evidence that human rights treaties systematically improve rights in signatories;⁶ and membership in a few treaties (especially the United Nations [UN] Convention Against Torture) actually seemed to be associated with deteriorating human rights conditions.⁷ Subsequent research, however, has generated evidence that treaties can enhance human rights under certain circumstances, which primarily depend on domestic politics or the nature of the rights in question.⁸ The debate is far from settled.

Human rights treaties, however, are not the only international institutions that aspire to promote human rights. A large number of IOs also aim—or claim—to do so, frequently by linking human rights to their other missions. For example, the Organization for Security and Co-operation in Europe (OSCE)—which is the world’s largest regional security organization, with 57 participating states—reports that respect for human rights and fundamental freedoms are key to the organization’s comprehensive security concept.⁹ Its Office for Democratic Institutions and Human Rights claims to be active in fostering human rights, tolerance, and non-discrimination, and its field operations provide members with information about and assistance in promoting rights. The

---

⁸ Conrad and Ritter 2013; Lupu 2013a, 2015; Powell and Staton 2009; Simmons 2009.
Organization for Economic Co-operation and Development (OECD) encourages policies that will improve the economic and social well-being of people around the world, integrating human rights into its development goals. The OECD has adopted formal guidelines to advance human rights-friendly corporate codes of conduct by multinationals operating in or from 42 countries.\textsuperscript{10}

It is not only IOs composed of Western democracies that articulate human rights goals. The Southern African Development Community (SADC) claims to work to ensure economic well-being, improve standards of living and quality of life, further freedom and social justice, and foster peace and security for members. The organization has also established formal principles and guidelines for promoting free and fair elections in its 15 member states based on “shared values and principles of democracy, the rule of law and respect for human rights.”\textsuperscript{11} The League of Arab States (now the Arab League) attempts to strengthen political, cultural, economic, and social programs among its members. The organization has a formal Permanent Committee on human rights (similar to the UN Human Rights Council), which is charged with establishing rules of cooperation on human rights issues among members, promoting implementation of human rights commitments, and monitoring implementation of the updated Arab Charter on Human Rights. The proliferation of IOs that claim to foster the promotion of human rights raises the question whether these organizations—many of which are more powerful in terms of resources and requirements than the universal UN human rights treaties—actually precipitate change in behavior. The answer, we argue, depends on the sovereignty costs they impose on members.

\begin{center}
Sovereignty Costs in International Organizations
\end{center}

\textsuperscript{10} See https://www.oecd.org/corporate/mne/1898155.pdf.

Human rights IOs exhibit variation in institutional design. Chief among these design features is the sovereignty costs imposed on members. Some create substantial incentives for members to comply with the organization, articulate clear and precise rules for members, and extensively monitor the actions of participating countries, whereas others do not. This variation in sovereignty costs is central to our argument.

To measure these costs, we rely on an index that is consistent with extant work on legalization and world politics developed by Hafner-Burton, Mansfield, and Pevehouse.\textsuperscript{12} This index taps each organizational rule’s precision, obligation, and delegation in the area of human rights.\textsuperscript{13}

First, Hafner-Burton, Mansfield, and Pevehouse code the precision of each human rights IO’s legal obligations. Greater precision reduces the room for signatory states to interpret rules in light of their own particular interests. Precise rules clarify which actions are acceptable and unacceptable, avoiding very general language that might permit behavior outside the spirit of the agreement.\textsuperscript{14} Their measure assigns higher sovereignty cost scores to organizations that make specific references to human rights norms, linking those norms to any of the organization’s other missions or specifying protection of human rights as their sole or central mission.\textsuperscript{15} Low sovereignty cost organizations are those whose founding documents only vaguely allude to the improvement of human rights without direct reference to international law. In the data that we

\textsuperscript{12} Hafner-Burton, Mansfield, and Pevehouse 2015.

\textsuperscript{13} Abbott and Snidal 2000; Abbott et al. 2000. We briefly review the Hafner-Burton, Mansfield, and Pevehouse coding scheme below. For more extensive detail, however, consult their study.

\textsuperscript{14} Abbott et al. 2000, 415.

\textsuperscript{15} Haas 1980.
analyze below, for example, the Inter-American Commission on Human Rights rates highly on the precision metric, while the International Organization for Migration ranks low because it “offer[s] little detail on which rights and what body of law or treaties are relevant.”16 Such imprecision furnishes members with substantial discretion in interpreting their obligations and provides few incentives for new members to actually improve human rights.

Second, Hafner-Burton, Mansfield, and Pevehouse code the degree of obligation for each human rights IO. High levels of obligation involve “unconditional pledges to be bound,” whereas low levels are “generated by mere recommendations or intent to remain bound by an agreement.”17 Sovereignty costs are higher when commitments are more binding and explicit.

The effects of obligations on sovereignty costs operate through several channels. To begin, there is variation in how formally organizations address human rights. Some maintain a formal apparatus to enforce human rights, while others have only informal mechanisms. In addition, issue linkage creates leverage with which organizations can incentivize member-states to comply with their obligations.18 By maintaining levers in other issue areas, human rights institutions can better encourage behavior consistent with human rights commitments.19 Finally, human rights IOs impose higher sovereignty costs if they have strict human rights or democracy requirements for membership (such as the European Union).20 As Hafner-Burton, Mansfield, and Pevehouse point out, it is not only European IOs that have such requirements. The SADC, for example, also

16 Hafner-Burton, Mansfield, and Pevehouse 2015, 6.

17 Hafner-Burton, Mansfield, and Pevehouse 2015, 6.

18 Haas 1980.

19 See Hafner-Burton 2005 for the example of trade and human rights.

explicitly requires members to promote equal opportunity and the political participation of their citizens. Many IOs, however, do not have explicit human rights membership criteria and therefore create relatively few explicit obligations for participating countries to improve rights.

Finally, Hafner-Burton, Mansfield, and Pevehouse consider the extent of delegation in each human rights IO. They operationalize delegation in terms of the degree to which IOs vest third parties—such as international adjudicators or non-governmental organizations (NGOs) or the IO itself—with the authority to perform dispute settlement, enforcement, and monitoring the behavior of members. Sovereignty costs are higher when IO-governed bodies (rather than those that are member-state governed) are granted more power over human rights. Sovereignty costs are also higher for those IOs that delegate more extensively to NGOs on human rights issues. Various studies of human rights have argued that delegation to NGOs increases the likelihood that states violating human right standards will be confronted over their non-compliant behavior.21 Further studies suggest that NGOs’ capacity to mobilize popular support and social movements increases the probability that a state will abide by its human rights commitments.22 Finally, Hafner-Burton, Mansfield, and Pevehouse also emphasize interaction among members and reporting of human rights in coding delegation. They note that some organizations “hold regular meetings or have the ability to call ad hoc meetings on human rights issues,” which increase sovereignty costs since states have chosen to delegate monitoring to an international institution.23

In sum, there is extensive variation in the design of human rights IOs. One key source of variation is the sovereignty costs imposed on members, which affect whether organizational claims

21 Keck and Sikkink 1998; Risse and Sikkink 1999.


to promote human rights produce incentives for acceding participants to improve their treatment of individuals. We now turn to testing the hypothesis that states joining human rights IOs are more likely to exhibit improvements in human rights behavior if the organization is marked by higher sovereignty costs than if these costs are lower.

Measuring Sovereignty Costs in Human Rights IOs

The most extensive data set on international organizations is compiled by the Correlates of War (COW) Project, but it does not code many organizations that are relevant to the study of human rights. For example, various human rights institutions are emanations, but the COW data excludes them, limiting its usefulness in a study such as ours.25

To fill this gap, Hafner-Burton, Mansfield, and Pevehouse identify and code state membership in organizations that explicitly claim to promote, advance, or enforce human rights in some capacity during the period from 1945 to 2000.26 They remove organizations that were largely non-governmental rather than inter-governmental and those that have rotating memberships.

While Hafner-Burton, Mansfield, and Pevehouse measure sovereignty costs in human rights IOs from 1945 to 2000, we expand their data through 2011 by performing a search using identical

24 Pevehouse, Nordstrom, McManus, and Jamison 2018.

25 For a thorough analysis of emanations, see Johnson 2014.

26 Hafner-Burton, Mansfield, and Pevehouse 2015, 8-9. Their sample was generated by “searching issues of the Yearbook of the Union of International Associations and examining every IO’s stated ‘Goals and Aims.’” They also “search[ed] previous electronic versions (on CD-ROMs) of the Yearbook using these criteria. We also made extensive use of the online Yearbook, searching for current and past (dead) organizations that fulfilled our criteria.”
criteria to theirs to identify human rights organizations formed from 2001 to 2011, finding none.

The resulting sample of 56 IOs is heterogeneous with respect to size, region, and scope of the organization’s overall mission. It includes large multilateral IOs such as the UN, large regional organizations such as the Organization of American States, and smaller regional groupings such as the SADC. The sample is fairly large and quite diverse, which allow us to make more significant generalizations about our findings. That the sample is heterogeneous with regard to types of international organizations also contributes to significant variation in the sovereignty costs generated by the organizations.

**Sovereignty Costs**

There is no scholarly consensus on how to measure the sovereignty costs that a human rights organization imposes on member-states. The first attempt was made by Cole, who gauges the extent of a state’s obligation to three UN human rights treaties by a state’s level of membership (signing a treaty incurs weaker obligation than ratifying) and a treaty’s degree of surveillance for monitoring compliance.\(^{27}\) While Cole’s effort is valuable, it does not address the role that international organizations play in furthering the cause of human rights. The second was by Hafner-Burton, Mansfield, and Pevehouse. As we discussed earlier, they code a sample of human rights IOs for ten factors which are linked to the concepts of precision, delegation, and obligation.

Two items tap precision: (1) is there a formal institution to focus on human rights and (2) are commitments grounded in specific and precise international laws? Four items capture delegation: (3) does the institution engage in formal reporting, (4) hold regular meetings concerning human rights, (5) have the ability to call ad hoc meetings on human rights, or (6) vest NGOs with authority

\(^{27}\) Cole 2009.
in the human rights arena? Four items measure obligation: (7) does the IO have a built-in enforcement mechanism, (8) are multiple issues discussed in the institution (facilitating issue linkage), and is membership conditional on a state being (9) democratic or (10) committed to promoting human rights? Each item is assigned a score of 0 or 1 and each IO’s sovereignty costs are calculated as the sum of the scores across these ten items. Clearly, this is not an exhaustive list of the institutional design characteristics that affect sovereignty costs, but these ten characteristics nonetheless play a crucial role in shaping such costs in the area of human rights.

The complete list of the 56 IOs in the sample and their corresponding sovereignty cost scores are shown in the online appendix. We follow Hafner-Burton, Mansfield, and Pevehouse in only coding sovereignty costs in the year that a state accedes to a human rights IO. In all other years, sovereignty costs are assigned a value of 0. For each organization, a higher score reflects greater precision, delegation, obligation, and therefore higher sovereignty costs.

**A Statistical Model of Human Rights**

To test our hypothesis about the influence of sovereignty costs on human rights outcomes, we estimate a series of statistical models. We begin with a sparse model, which we then populate with additional control variables:

\[
\Delta \text{Human Rights Behavior}_{t+1} = \gamma_0 + \gamma_1 \text{Sovereignty Costs} + \varepsilon_1 \quad [1]
\]

\[
\Delta \text{Human Rights Behavior}_{t+1} = \gamma_0 + \gamma_1 \text{Sovereignty Costs} + \gamma_2 \text{Human Rights Behavior}_t + \varepsilon_1 \quad [2]
\]

\[
\Delta \text{Human Rights Behavior}_{t+1} = \gamma_0 + \gamma_1 \text{Sovereignty Costs} + \gamma_2 \text{Human Rights Behavior}_t + \gamma_3 \text{Polity} + \gamma_4 \Delta \text{Polity} + \gamma_5 \text{Population} + \gamma_6 \text{GDP per capita} + \varepsilon_1
\]

---

28 Hafner-Burton, Mansfield, and Pevehouse, 10-12.
\[
\gamma_{\text{Former Communist}} + \gamma_{\#\text{Human Rights Organizations}} + \epsilon_i
\]  

In our study, the unit of analysis is the country-year. To construct our dependent variable, we use Fariss’ latent variable model of human rights behavior to generate scores for a large cross-section of states between 1950 and 2012. Fariss points out that latent variable estimation accounts for changes in a country’s human rights behavior, but also adjusts for changing standards of accountability over time. Consequently, ratings of countries may change because of shifts in state behavior, but also because of shifting human rights standards adopted by the international community or others who evaluate this behavior (e.g., the United States Department of State). Due to concerns about the possibility of endogeneity (which we discuss more fully below), the slow-moving nature of this variable, and the attendant issues that could arise with autocorrelation, we analyze the first difference in the variable. Thus, \( \Delta \text{Human Rights Behavior}_{t+1} \) is the annual change in human rights behavior for a given state \( i \) from year \( t \) to year \( t+1 \).

Our key independent variable—\textit{Sovereignty Costs}—is the sovereignty costs associated with any human rights IO that state \( i \) accedes to in year \( t \). If state \( i \) joins multiple human rights IO in a given year, this variable is coded as the maximum sovereignty cost score among the organizations that it entered. If a state does not join a human rights institution in year \( t \), its sovereignty cost score for that year is zero.

---

29 Fariss 2014.

30 The data are available at cfarris.com.

31 This includes IOs in which state \( i \) is an associate member. We exclude cases where the state is an observer.

32 It is possible that a state could join an institution with a sovereignty cost score of zero, thus conflating not joining a human rights institution with joining a zero-cost institution. This is very
Our next model adds a measure of each state’s human rights behavior in year $t$. Doing so accounts for the possibility of floor or ceiling effects in such behavior. If, for example, states already exhibit good behavior, there will be less room for further improvements and hence a reduced likelihood of observing increases in the dependent variable.

Our third model includes a number of additional controls. First, $Polity$ is state $i$’s regime type in year $t$, based on the 21-point index drawn from the Polity IV data set that ranges from 10 for the most democratic states to -10 for the most autocratic ones.$^{33}$ Past research has shown that regime type is a key determinant of both human rights behavior as well as membership in both international institutions covering human rights and high sovereignty cost institutions.$^{34}$ Second, $\Delta Polity$ is the difference in state $i$’s regime type from year $t-1$ to year $t$. Previous studies suggest that states undergoing regime change in a democratic direction are more likely to join high sovereignty cost human rights institutions.$^{35}$ States undergoing democratic regime change are also more likely to experience improvements in their human rights behavior.$^{36}$

---

rare in the data. Below, we re-estimate our main models using only cases where any human rights institution was joined as a robustness check.


$^{34}$ On human rights, see Landman 2005. On institutions more generally, see Mansfield and Pevehouse 2006. On sovereignty costs, see Hafner-Burton, Mansfield, and Pevehouse 2015.

$^{35}$ Hafner-Burton, Mansfield, and Pevehouse 2015.

$^{36}$ Substituting indicator variables for democratization and autocratization yields very similar estimates to those reported with the differenced regime type variable. On the construction of these indicators, see Mansfield and Snyder 2005.
Third, Population and GDP per capita, which are measured in year \( t \), are introduced since both variables typically appear in models of human rights compliance.\(^{37}\) Fourth, Former Communist equals 1 for states that were communist at some point during the post-World War II era, beginning in the year after the communist government fell, 0 otherwise. Many states that had a communist government expressed a desire to join Western IOs after the Cold War, including those with a human rights mandate. We need to ensure that these cases are not driving any observed relationship between human rights behavior and sovereignty costs in newly joined human rights institutions.

Finally, because entry into a given human rights IO is undoubtedly related to the number of organizations to which state \( i \) is already a member, we include \#Human Rights Organizations. As a state participates in a growing number of human rights IOs, the marginal benefit of joining another one may decline and as the number of human rights IOs that it does not belong to falls, it reduces the number that it could enter. Alternatively, states that participate in a large number of human rights organizations may be “joiners,” predisposed to accede to as many of these institutions as possible. Those that belong to few human rights organizations may have an aversion to joining more.

In each model, we include fixed effects for countries and years. Doing so accounts for unobserved heterogeneity across counties that might influence changes in their human rights behavior. It also accounts for factors that occur at the global level and that could influence the behavior of various countries at a given point in time.

Results

All of the models are estimated using ordinary least squares. Table 1 presents the initial results. In each case, the estimated coefficient of *Sovereignty Costs* is positive and statistically significant, indicating that joining higher sovereignty cost IOs leads to greater improvements in the human rights outcomes of new members than joining lower cost IOs. Moreover, including measures of current human rights behavior, regime type, population, GDP per capita, and whether a state was formerly communist has little bearing on the estimated coefficient of *Sovereignty Costs*.

Nonetheless, some of these variables influence changes in human rights behavior. First, the coefficient estimate of *Human Rights Behavior* is negative and statistically significant, which is no doubt due to a ceiling effect. States that exhibit more respect for human rights have less room to further improve their behavior than states with a poorer rights record. Second, as shown in column 3, both *Polity* and Δ*Polity* have positive and statistically significant effects on human rights behavior. Consistent with prior work, these results indicate that more democratic states and those becoming more democratic tend to display greater respect for human rights than other countries.38

Third, *#Human Rights Organizations* has a negative influence on human rights. This is an important finding for two reasons. That its inclusion in our model does not influence the estimate of *Sovereignty Costs* indicates that both the number of organizations and their institutional design influence human rights performance. Moreover, these results suggest that there are diminishing returns from joining human rights organizations. Finally, more populous and wealthier countries

38 Note that results based on Cheibub, Ghandi, and Vreeland’s (2010) regime type data and the V-Dem Project’s polyarchy index (Coppedge et al. 2017) yield very similar results to those presented here. Also note that there is no evidence of an interaction effect between regime type and sovereignty costs. These results are available in the online appendix.
Table 1. The determinants of change in human rights behavior

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sovereignty Costs</td>
<td>0.002**</td>
<td>0.002*</td>
<td>0.002*</td>
<td>0.008*</td>
<td>0.002*</td>
</tr>
<tr>
<td></td>
<td>(0.001)</td>
<td>(0.001)</td>
<td>(0.001)</td>
<td>(0.002)</td>
<td>(0.001)</td>
</tr>
<tr>
<td>Human Rights Behavior</td>
<td>-0.028**</td>
<td>-0.033**</td>
<td>-0.038**</td>
<td>-0.033**</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.002)</td>
<td>(0.003)</td>
<td>(0.008)</td>
<td>(0.003)</td>
<td></td>
</tr>
<tr>
<td>Polity</td>
<td>0.002**</td>
<td>0.000</td>
<td>0.002**</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.000)</td>
<td>(0.001)</td>
<td>(0.000)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ΔPolity</td>
<td>0.006**</td>
<td>0.012**</td>
<td>0.006**</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.001)</td>
<td>(0.003)</td>
<td>(0.001)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GDP per capita (ln)</td>
<td>-0.015**</td>
<td>-0.029*</td>
<td>-0.015**</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.005)</td>
<td>(0.015)</td>
<td>(0.005)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Population (ln)</td>
<td>-0.078**</td>
<td>-0.082*</td>
<td>-0.078**</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.012)</td>
<td>(0.036)</td>
<td>(0.012)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Former Communist</td>
<td>0.004</td>
<td>0.039</td>
<td>0.005</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.016)</td>
<td>(0.045)</td>
<td>(0.016)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>#HR Organizations</td>
<td>-0.004**</td>
<td>-0.007*</td>
<td>-0.004**</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.001)</td>
<td>(0.002)</td>
<td>(0.001)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Regional Sov. Costs</td>
<td></td>
<td></td>
<td></td>
<td>-0.000</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(0.000)</td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>0.007</td>
<td>0.012</td>
<td>0.267**</td>
<td>0.324**</td>
<td>0.267**</td>
</tr>
<tr>
<td></td>
<td>(0.015)</td>
<td>(0.015)</td>
<td>(0.051)</td>
<td>(0.150)</td>
<td>(0.051)</td>
</tr>
<tr>
<td>Observations</td>
<td>9,520</td>
<td>9,520</td>
<td>7,284</td>
<td>1,132</td>
<td>7,284</td>
</tr>
<tr>
<td>Number of countries</td>
<td>202</td>
<td>202</td>
<td>155</td>
<td>155</td>
<td>155</td>
</tr>
</tbody>
</table>

Note: Entries are ordinary least squares regression estimates with standard errors in parentheses. Two-tailed tests of statistical significance are conducted for all coefficient estimates. Statistical significance is indicated as follows: ** p < 0.01, * p < 0.05.
experience fewer improvements in human rights behavior than their smaller and poorer counterparts, as the estimated coefficients of *Population* and *GDP per capita* are negative and statistically significant. However, whether a country had a communist government has little bearing on change in its human rights behavior.

Our fourth set of estimates in Table 1 addresses the effect of our decision to code years when a state did not join any human rights IO as zero. When we re-estimate the model after eliminating each observation where a state did not accede to an IO, the results in column 4 show that the estimated coefficient of *Sovereignty Costs* remains positive and statistically significant. Furthermore, none of the other independent variables change sign and only one experiences a change in statistical significance. Consequently, the decision to code non-joining years as zero has little impact on our results.

Next, it is useful to account for regional patterns of human rights and institutions since extant scholarship on treaties suggests that regional diffusion can affect respect for human rights, often through regional institutions themselves. To address this issue, we include a measure of the annual sovereignty cost score of each geographic region. The estimates of the model with this additional variable are reported in column 5 of Table 1 and provide no evidence that the measure of regional sovereignty costs is either statistically significant or affects the estimate of *Sovereignty Costs*.

As an additional precaution, we estimate a model controlling for the act of joining a human rights IO, regardless of its level of sovereignty costs. To ensure that the observed effects of sovereignty costs do not simply reflect the influence of joining a human rights IO, we re-estimate the model after adding an indicator variable coded as 1 if a state joins a human rights IO in year $t$.

---

39 Greenhill 2015.

40 We use the COW System Membership data to code geographic regions (Correlates of War 2017).
and 0 otherwise. This variable does not achieve statistical significance; nor does it affect the statistical significance or sign of the coefficient of Sovereignty Costs. Hence, changes in human rights behavior are being driven by sovereignty costs and not by whether or not a state joins any human rights IO.

Finally, some readers may be concerned about the aggregate nature of the Hafner-Burton, Mansfield, and Pevehouse sovereignty cost index. For example, is it possible that only one of the three facets of sovereignty costs included in the index—precision, obligation, or delegation—is driving the estimates? To assess this possibility, we re-estimate the model after replacing the aggregate measure of sovereignty costs with the respective scores for precision, obligation, and delegation. The estimated coefficient of each of these three variables, when introduced individually, is positive and statistically significant, indicating that all three features of sovereignty costs influence a state’s human rights behavior.41

**Addressing the Possibility of Endogeneity**

As with any study attempting to assess the effects of institutions, it is important to address whether our results are threatened by endogeneity. Various observers have pointed out that international institutions should be treated as endogenous if the very behavior they regulate is a key factor shaping which states joining them.42 In our case, it is possible that states with better (worse) human rights records select themselves into higher (lower) sovereignty cost organizations. We have taken several steps to avoid this possibility in order to make the cleanest possible inference about

---

41 We introduce each component individually because they are highly correlated, with bivariate correlations between the three variables ranging from 0.78 to 0.89. These estimates, as well as those discussed in the previous paragraph, are available in the online appendix.

the effects of joining human rights IO.

First, it is worth noting that the classic argument about non-random selection into IOs suggests that only well-behaved states would join demanding institutions.43 Yet, in the human rights arena, various studies have found that such states often do not join international regimes and that poorly-behaved states frequently do enter them.44 Thus, if selection creates a bias, its direction is theoretically indeterminate.

Second, we introduce a variable to control for the underlying preferences of states to join international institutions. In this way, we can account for a heretofore unobservable factor that could contribute to endogeneity. Lupu has developed a measure to capture “treaty commitment preferences,” which is an indicator of a state’s underlying propensity to enter human rights treaties.45 We add this variable (which we label Propensity Measure) to the model to determine if it influences the observed effect of Sovereignty Costs. If Sovereignty Costs is beset by non-random selection stemming from a state’s desire or decision to join a human rights IO, rather than the actual effect of the institution’s design, then including this propensity measure should erode the size and statistical significance of Sovereignty Costs since the measure accounts for unobservable traits that prompt states to join treaties. As shown in the first column of Table 2, this is not the case. Lupu’s measure is not statistically significant; nor does including it erode the significance of Sovereignty Costs.

43 Downs, Rocke, and Barsoom 1996.


45 Lupu 2013b.
Table 2. Tests addressing endogeneity and intertemporal effects of IOs on changes in human rights behavior

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>ΔHR Behavior_{t+1}</td>
<td>ΔSovereignty Costs_{t+1}</td>
<td>ΔHR Behavior_{t+1}</td>
<td>ΔHR Behavior_{t+1}</td>
<td>ΔHR Behavior_{t+1}</td>
</tr>
<tr>
<td>Sovereignty Costs_{t}</td>
<td>0.002*</td>
<td>-0.997**</td>
<td>0.002*</td>
<td>0.198**</td>
<td>0.002*</td>
</tr>
<tr>
<td></td>
<td>(0.001)</td>
<td>(0.011)</td>
<td>(0.001)</td>
<td>(0.062)</td>
<td>(0.001)</td>
</tr>
<tr>
<td>Sovereignty Costs_{t-1}</td>
<td>0.002*</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.001)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sovereignty Costs_{t-2}</td>
<td>0.002*</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.001)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sovereignty Costs_{t-3}</td>
<td>0.002*</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.001)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sovereignty Costs_{t-4}</td>
<td>0.002*</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.001)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sovereignty Costs_{t-5}</td>
<td>0.001</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.001)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sovereignty Costs_{t-6}</td>
<td>0.001</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.001)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sovereignty Costs_{t-7}</td>
<td>-0.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.001)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sovereignty Costs_{t-8}</td>
<td>-0.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.001)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sovereignty Costs_{t-9}</td>
<td>0.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.001)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Human Rights Behavior</td>
<td>-0.032**</td>
<td>0.049</td>
<td>-0.033**</td>
<td>-0.033**</td>
<td>-0.034**</td>
</tr>
<tr>
<td></td>
<td>(0.003)</td>
<td>(0.039)</td>
<td>(0.003)</td>
<td>(0.003)</td>
<td>(0.003)</td>
</tr>
<tr>
<td>Polity</td>
<td>0.001*</td>
<td>0.027**</td>
<td>0.001**</td>
<td>0.000</td>
<td>0.002**</td>
</tr>
<tr>
<td></td>
<td>(0.000)</td>
<td>(0.006)</td>
<td>(0.000)</td>
<td>(0.001)</td>
<td>(0.000)</td>
</tr>
<tr>
<td>ΔPolity</td>
<td>0.006**</td>
<td>0.023</td>
<td>0.006**</td>
<td>0.005**</td>
<td>0.006**</td>
</tr>
<tr>
<td></td>
<td>(0.001)</td>
<td>(0.013)</td>
<td>(0.001)</td>
<td>(0.001)</td>
<td>(0.001)</td>
</tr>
<tr>
<td>GDP per capita (ln)</td>
<td>-0.017**</td>
<td>0.043</td>
<td>-0.015**</td>
<td>-0.016**</td>
<td>-0.017**</td>
</tr>
<tr>
<td></td>
<td>(0.005)</td>
<td>(0.065)</td>
<td>(0.005)</td>
<td>(0.005)</td>
<td>(0.006)</td>
</tr>
<tr>
<td>Population (ln)</td>
<td>-0.061**</td>
<td>-0.481**</td>
<td>-0.077**</td>
<td>-0.078**</td>
<td>-0.091**</td>
</tr>
<tr>
<td></td>
<td>(0.014)</td>
<td>(0.153)</td>
<td>(0.012)</td>
<td>(0.013)</td>
<td>(0.014)</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>--------------------------------</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>Former Communist</td>
<td>0.005</td>
<td>1.710**</td>
<td>0.001</td>
<td>-0.028</td>
<td>-0.004</td>
</tr>
<tr>
<td></td>
<td>(0.017)</td>
<td>(0.219)</td>
<td>(0.016)</td>
<td>(0.019)</td>
<td>(0.017)</td>
</tr>
<tr>
<td>#HR Organizations</td>
<td>-0.004**</td>
<td>-0.168**</td>
<td>-0.004**</td>
<td>-0.002</td>
<td>-0.005**</td>
</tr>
<tr>
<td></td>
<td>(0.001)</td>
<td>(0.012)</td>
<td>(0.001)</td>
<td>(0.001)</td>
<td>(0.001)</td>
</tr>
<tr>
<td>Propensity Measure</td>
<td>0.011</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.010)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>0.259**</td>
<td>4.273**</td>
<td>0.260**</td>
<td>-0.093</td>
<td>0.297**</td>
</tr>
<tr>
<td></td>
<td>(0.057)</td>
<td>(0.673)</td>
<td>(0.051)</td>
<td>(0.113)</td>
<td>(0.058)</td>
</tr>
<tr>
<td>Observations</td>
<td>6,592</td>
<td>7,434</td>
<td>7,284</td>
<td>6,972</td>
<td>6,594</td>
</tr>
<tr>
<td>Number of countries</td>
<td>151</td>
<td>155</td>
<td>155</td>
<td>154</td>
<td>154</td>
</tr>
</tbody>
</table>

Note: Entries are ordinary least squares estimates with standard errors in parentheses. Two-tailed tests of statistical significance are conducted for all estimates. Statistical significance is indicated as: ** $p < 0.01$, * $p < 0.05$. In column 3, the dependent variable is adjusted using the sequential-G estimation procedure and the standard error of Sovereignty Costs is computed via bootstrap. In column 4, the Sovereignty Cost measure is replaced by an instrument—the predicted value of Sovereignty Costs based on the model in Hafner-Burton, Mansfield, and Pevehouse 2015. The standard error of Sovereignty Costs is computed via bootstrap.
Previous empirical work has found little evidence that: (1) human rights behavior is correlated with the decision to join human rights organizations or (2) conditional on an affirmative decision to join, that human rights behavior influences decisions about whether to enter one marked by higher sovereignty costs. Nonetheless, we address this issue since one implication of arguments about endogenous institutions and endogenous cooperation is that states with better human rights behavior should join higher sovereignty cost organizations. More specifically, we assess the impact of human rights behavior in year $t$ on change in the value of Sovereignty Costs between years $t$ and $t + 1$. As shown in the second column of Table 2, the coefficient estimate of Human Rights Behavior is not statistically significant.

Next, we re-estimate our model to address concerns raised by Blackwell and Glynn that lagging explanatory variables to avoid endogeneity can create biased estimates. We use their suggested Sequential-G estimation technique, which adjusts for the influence of prior effects of explanatory variables by removing these effects on the dependent variable before re-estimating the model. As shown in column 3 of Table 2, we obtain nearly identical results using this technique. Taken together, these tests provide no indication our results are undermined by endogeneity.

Finally, we construct an instrument for Sovereignty Costs using the predicted values from a

---

46 Hathaway 2007 finds that only some states may be deterred from joining human rights commitments due to past behavior. Hollyer and Rosendorff 2011b, in their first stage models of joining the Convention Against Torture, find mixed results on the effect of existing levels of torture. Most studies examining the effect of past human rights behavior are limited to treaties. Hafner-Burton, Mansfield, and Pevehouse 2015, based on a broader sample of treaties and IOs, find no correlation between previous human rights behavior and accession to high sovereignty cost IOs.

47 Blackwell and Glynn 2018.
statistical model of sovereignty costs in a study that we discussed earlier conducted by Hafner-Burton, Mansfield, and Pevehouse.\textsuperscript{48} That model jointly predicted the probability of acceding to a human rights IO and the associated sovereignty costs in each country-year.\textsuperscript{49} We replace our current measure, \textit{Sovereignty Costs}, with the predicted level of sovereignty costs for each country-year from that model. The sovereignty cost model contains a measure of hegemony (defined as the portion of global military capabilities controlled by the largest state—in our case, the U.S.), which serves as an exclusion restriction for our current model. Previous research has found that hegemony promotes the establishment of human rights regimes—indeed, the U.S. has been at the forefront of negotiations to create (if not always join) both regional and international human rights treaties and courts.\textsuperscript{50} American “exceptionalism,” however, suggests that hegemony is likely to be unrelated to human rights behavior given the U.S.’s spotty record of human rights enforcement throughout the globe.\textsuperscript{51} Thus, hegemony’s influence on human rights outcomes is most likely to occur through its effect on the establishment and strength of human rights institutions. Column 4 of Table 2 shows the results of this model. The coefficient estimate of the instrument (with a bootstrapped standard error) is positive and statistically significant. This suggests that the predicted value of \textit{Sovereignty Costs}—our instrumental variable—is strongly related to human rights behavior.

\begin{footnotesize}
\begin{enumerate}
\item Hafner-Burton, Mansfield, and Pevehouse 2015.
\item Hafner-Burton, Mansfield, and Pevehouse relied on a zero-inflated negative binomial model. We use our temporally expanded data to re-estimate their model to recover predicted values. We are able to recover their original result with our expanded data.
\item Donnelly 1986; Moravcsik 2000.
\item Ignatieff 2005; Moravcsik 2005.
\end{enumerate}
\end{footnotesize}
Intertemporal Effects

Our statistical models have estimated the short-run effects of human rights institutions in the year after accession. But joining human rights organizations may have longer-term effects as well. Although this point has been made repeatedly in the theoretical literature on human rights institutions, it is rarely assessed empirically.\textsuperscript{52} We therefore re-estimate our model after adding nine consecutive one-year lags of sovereignty costs to capture any over-time effects in changing human rights behavior.\textsuperscript{53}

The estimates presented in column 5 of Table 2 indicate that the estimated coefficient of our original measure of Sovereignty Costs (at time $t$) remains nearly identical to that based on our earlier results. The next four lags, from $t-1$ through $t-4$, are also positive and statistically significant. Thereafter, sovereignty costs appear to have little influence. That joining high sovereignty cost organizations continues to improve human rights for about five years, after which the behavior of governments seems to plateau, may reflect the time needed for the human rights provisions of IOs to be integrated into a country’s legal system or to mobilize domestic groups with an interest in compliance with these provisions.\textsuperscript{54} Regardless, it is notable that such IOs stimulate both short-term and medium-term improvements in human rights.

\textsuperscript{52} See Chayes and Chayes 1993 on the over-time effects of human rights institutions.

\textsuperscript{53} This type of model is akin to an autoregressive distributed lag (ADL) model. We introduce these lags into the model as linear variables, whereas most traditional ADL models create transformations of the lags (usually polynomials) to avoid multicollinearity. Since multicollinearity is not a concern in our model, we choose to keep a simplified version of the model. Transforming the lags as polynomials yields very similar results.

\textsuperscript{54} Simmons 2009.
Finally, it is important to recognize that the influence of Sovereignty Costs on human rights behavior is substantively important, as well as statistically significant. Based on the results in column 3 of Table 1, predicted human rights behavior improves from about 0.016 to 0.029 in the year after joining an IO with the maximum sovereignty cost score of 10. While this might seem to be a modest rise, it is substantively meaningful since ΔHuman Rights Behavior ranges from -0.88 to 0.87, with a mean of 0.01. Moreover, it is widely argued that a democratic regime type promotes human rights, and the size of this predicted improvement is about the same as that associated with a change from an anocracy (a Polity score of 0) to a democracy (the maximum Polity score of 10).

But focusing on the first year of IO membership understates the substantive importance of Sovereignty Costs. Based on the results in column 5 of Table 2, the cumulative effect over the first five years of membership of acceding to an IO with the maximum sovereignty cost score is to improve the predicted human rights behavior by roughly 0.11, which is sizable.

Conclusions

The rise of international human rights institutions has spawned a vigorous debate in the fields of international relations and international law. Almost all of the research conducted on this issue addresses whether participation in a handful of human rights treaties leads to improvements in human rights. Much of that debate now revolves around the domestic scope conditions necessary to generate state compliance with treaty obligations. We engage the debate from a different angle by exploring whether variation in the design of international organizations that claim to promote human rights influences the human rights performance of member-states. A large and growing number of IOs claim to incorporate the promotion of human rights into their formalized goals. These organizations often have substantial power. Less clear, however, is whether that rhetoric actually prompts members to improve their human rights records. By using a new data set of
organizations with human rights mandates, and mapping variation in the sovereignty costs associated with participation, we examined whether and how these IOs are able to improve human rights on the ground.

We find significant support for the argument that membership in IOs marked by higher sovereignty costs tends to stimulate improvements in human rights. Moreover, these improvements continue for about five years after accession. This finding is robust to the inclusion of a series of control variables and alternative statistical models, and we show that the observable implications of critiques concerning the selection of states into certain types of human rights IOs are not supported. There is no evidence that states strategically select into low or high sovereignty cost organizations.

Our findings bear on debates about whether human rights instruments are effective, as well as debates about the optimal design of those regimes. Furthermore, they speak directly to the wisdom of linking human rights to other, often more powerful institutions outside the body of international treaties designed to promote the issue. Unless IOs impose significant sovereignty costs, discussions of human rights remain largely empty rhetoric. When they do impose such costs, issue linkage outside the treaty regime can be effective.
References


