The Strength of Weak Review: 
National Courts, Interpretive Canons, and Human Rights Treaties

Yonatan Lupu, George Washington University
Pierre-Hugues Verdier, University of Virginia
Mila Versteeg, University of Virginia

Abstract

What is the role of national courts in international judicialization? National courts are often key to the enforcement of human rights treaties. How do they do so? We focus on two mechanisms: direct effect, under which courts can apply treaties directly and set aside inconsistent domestic laws; and a canon of interpretation, under which courts interpret domestic laws in conformity with treaties. We argue that the legally weaker canon mechanism is more politically effective because it minimizes opposition to court rulings. Drawing on new data, we find that the canon significantly increases the extent to which ICCPR ratification reduces abuses of the freedoms of speech, association, and religion. For the freedoms of religion and association, this effect is significantly greater than that of the direct effect mechanism. We also contribute to our understanding of the relationship between the adjudication, compliance, and feedback phases of judicialization.
1. Introduction

What is the role of national courts in international judicialization? Because international law is often incorporated into national law, national courts hear cases about alleged violations of international commitments. To what extent does such litigation promote the enforcement of international law? What are the legal mechanisms national judges use to enforce international law, and what is the relationship between those legal mechanisms and political constraints?

We address these questions in this paper focusing on how national courts enforce international human rights treaties. Existing research has identified several mechanisms by which domestic political actors and institutions can make human rights commitments effective. By mobilizing support for human rights, actors can exert pressure on governments to honor their human rights commitments. Actors can also rely on treaties to strengthen normative arguments regarding respect for human rights. By relying in part on the information politics activated by other actors, a sufficiently strong legislative opposition can increase the costs to the government of violating its treaty commitments.

Many scholars have focused on domestic courts as key actors in making human rights treaties more effective. Rights-claiming litigation under international human rights treaties is one of the key indicators of the judicialization of international relations. Domestic courts can overturn laws and impose penalties, which can both punish and deter human rights violations. Of course, domestic courts are subject to both political and legal constraints. Politically, if a domestic court is not sufficiently independent or effective, it may not have the power or authority to deter or punish violations by executives and legislatures. Legal constraints, such as the laws of evidence, also affect the ability of courts to enforce human rights treaties. The cost of obtaining legally admissible evidence with respect to killings and torture can be prohibitively high, making even independent domestic courts less effective as an enforcement mechanism against these violations.

While we have learned much regarding the conditions under which domestic courts can make human rights treaties more effective, the literature has not systematically examined how they do so. We build on this literature by asking the following question: by which mechanism(s) can domestic courts apply international human rights treaties, and how does the availability of these mechanisms affect human rights outcomes? The assumption in the existing literature appears to be that, once a state ratifies a human rights treaty, domestic courts are empowered to review government action and provide remedies for violations of the rights protected by the treaty. However, the application of treaties by domestic courts is mediated by legal doctrines

---

2 Simmons 2009.
3 Keck and Sikkink 1998.
4 Lupu 2015.
5 Alter, Hafner-Burton, and Helfer (2016).
6 Keith 2002; Powell and Staton 2009.
7 Lupu 2013a.
that govern the status of treaties in the domestic legal system. These legal doctrines have political consequences that have not previously been analyzed in the literature on the effects of human rights treaties. While they can empower courts to enforce international human rights commitments, they often limit their ability to do so directly.

We focus on two legal mechanisms that domestic courts may use to give effect to international human rights treaties. A growing number of states allow their courts to apply treaties directly, and provide that treaties trump ordinary legislation. In those states, treaties become quasi-constitutional and courts can use them to set aside domestic laws. We refer to this as the *direct effect* mechanism. Second, many states recognize doctrines under which domestic law should be interpreted in conformity with treaties. These doctrines allow courts to modify how domestic laws are interpreted without formally invalidating them. We refer to this as the *canon of interpretation* mechanism.

We argue that the extent to which courts can use these mechanisms to enforce human rights treaties depends on the possibility of political backlash against their decisions. Even independent judiciaries in countries with strong respect for the rule of law are attentive to the possibility of backlash, which threatens their legitimacy. The use of the direct effect mechanism is more likely to lead to backlash because it is more likely to be viewed as a political or abusive use of judicial power. By contrast, when courts use the canon of interpretation mechanism, their decisions are more likely to be perceived as legitimate by other government actors and the public, because interpreting the law is perceived as being a core judicial, rather than political, function. Thus, while at first glance direct effect appears to be the stronger legal mechanism, we generally expect the seemingly weaker mechanism of the canon of interpretation to be a more politically effective tool for domestic courts to enforce human rights treaties.

To test this argument, we focus on the effects of International Covenant on Civil and Political Rights (ICCPR) ratification. Domestic courts regularly apply ICCPR provisions via the mechanisms we highlight. Drawing on new data on the status of international law in domestic legal systems, we measure the extent to which each mechanism is available to a country’s domestic courts. Our empirical models focus on abuses of civil rights because those are the types of human rights abuses with respect to which domestic courts are likely to have effects.

Our findings support our argument. We find that the availability of the canon of interpretation mechanism to domestic courts significantly increases the extent to which ICCPR ratification reduces abuses of the freedoms of speech, association, and religion. Just as importantly, with respect to the freedoms of religion and association, our results indicate that the extent to which the canon of interpretation mechanism increases the effects of ICCPR ratification is significantly greater than the extent to which the direct effect mechanism does so. In other words, the canon of interpretation mechanism, which is a legally weaker form of review, is more politically effective. By contrast, with respect to freedom of speech, the availability of either legal mechanism appears to increase the effects of the ICCPR, and one mechanism does

---

8 Sandholtz 2012.
9 Vanberg 2005; Carrubba 2009.
10 Lupu 2013a.
not appear to do so more than the other. In section 5.1, we discuss why our results with respect to free speech may be driven by the fact that right is defined more precisely.

This paper makes several contributions to the literature. First, ours is the first study to systematically analyze the political consequences of the legal mechanisms by which domestic courts can enforce human rights treaties. In so doing, we build on and extend existing work in this area by demonstrating the conditions under which courts in countries that have joined human rights treaties can reduce human rights abuses. Second, we contribute to our knowledge about the role of national courts in international judicialization. Our argument indicates that the adjudication phase of judicialization politics is closely wound with the compliance and feedback phases. Judges, concerned with the possibility of backlash in the feedback phase, may strategically alter their behavior in the adjudication phase, which in turn affects the extent of compliance. Our argument also suggests that in order to understand the relationship between the phases of judicialization, we should consider legal adjudication mechanisms and the political consequences of those mechanisms. Finally, we provide new empirical results regarding the conditions under which human rights treaty commitment can lead to improved human rights practices.

The rest of this paper proceeds as follows. In section 2, we briefly discuss the key arguments and empirical results on which our argument builds. Section 3 describes our theory regarding the mechanisms by which domestic courts can enforce international human rights law. In section 4, we describe the research design we use to test our hypotheses. Section 5 presents our results. Section 6 offers conclusions and implications for future research.

2. Human Rights Abuses, International Treaties, and Domestic Courts

2.1. Constraining Repression

Domestic opposition groups threaten leaders’ ability to stay in power. Faced with such threats, leaders often respond with repression, although alternative responses can include cooption, accommodation, persuasion, and neglect. Repression deters potential challengers to the state and can raise the cost of mobilizing to challenge the state by disrupting the dissenting group, cutting off communication, making assembly more difficult, and restricting access to resources. Yet governments that abuse human rights also incur costs, including a loss of legitimacy, potentially generating new grievances against the government, and a reduction in coercive capacity. Governments have a repertoire of possible strategies of repression they can choose from, including indiscriminate mass atrocities, targeted killings and torture, limitations on the rights of groups to freely associate or practice their religion, restrictions on the media, and other limits on speech.

---

Can human rights treaties prevent governments from using repression? Recent years have witnessed a healthy debate over whether the international human rights regime has failed, succeeded, or otherwise yielded benefits worthy of its costs. Several prominent books have pronounced the endtimes or twilight of international human rights law. Others have examined the conditions and mechanisms through which human rights treaty ratification may reduce repression, often emphasizing the role of domestic institutions and political actors. Several theoretical and empirical studies have demonstrated that human rights treaty ratification can deter repression by changing leaders' incentives, although these effects are conditional and cannot be expected to prevent all human rights abuses. Others who believe in the power of human rights law focus on its persuasive and legitimizing power to change social norms, identities, and preferences, especially through the work of advocacy groups.

Among the domestic institutions that may make human rights treaties effective, domestic courts have long been understood to play a central role. Both the Universal Declaration of Human Rights and the ICCPR note that an independent judiciary is crucial to maintaining respect for human rights. In countries that have joined human rights treaties, a sufficiently powerful domestic judiciary may serve an enforcement function. Such courts can prosecute government actors or strike down laws that violate human rights, thus making governments less likely to carry out abuses.

2.2 Domestic Courts and Their Audiences

While courts are crucial for making treaty rights effective, they are also subject to important constraints. Some of these constraints arise from the nature of the judicial process. For example, courts can only act on the basis of legally admissible evidence that is sufficiently strong to meet the relevant standards of proof. With respect to abuses of personal integrity rights such as killings or torture, the government is often able to prevent courts from obtaining such evidence. As a result, even powerful, independent domestic judiciaries are likely to serve as enforcers of human rights treaties primarily with respect to abuses of civil rights, such as the freedoms of speech, association, and religion.

Even in the area of civil rights, courts face important political constraints. Indeed, the judiciary ultimately depends on the executive branch to enforce its decisions; yet, the executive may be the primary source of human rights violations. The literature has proposed at least two explanations for why the political branches comply with judicial decisions.

Some focus on the functional utility of courts, and specifically their ability to clarify law and to provide focal points for coordination among the political branches. The judiciary is valued as a neutral arbiter that helps political actors to coordinate upon a set of rules by which they

---

16 Simmons 2009; Sikkink 2011.
17 Keith 2002; Powell and Staton 2009.
18 Lupu 2013a.
conduct their political affairs.\textsuperscript{19} While the initial task of coordinating government falls upon the constitution itself, the judiciary can further aid coordination by clarifying the rules and announcing when political actors have overstepped their powers. Because courts can bring important coordination benefits, political actors will value having an independent court that can interpret the constitution. It is for this same reason that they comply with judicial rulings, even unfavorable ones.\textsuperscript{20}

Others argue that courts owe compliance to their legitimacy.\textsuperscript{21} Courts can draw on what David Easton called “diffuse support,” that is “a reservoir of favorable attitudes or good will” towards the institution overall, “that helps members to accept or tolerate outputs to which they are opposed.”\textsuperscript{22} Diffuse support, or legitimacy, is distinct from “specific support,” that is “the satisfactions that members of a system feel they obtain” from specific decisions and outputs.\textsuperscript{23} High levels of diffuse support can mitigate the effects of dissatisfaction with an unpopular opinion that enjoys low levels of specific support.\textsuperscript{24} Thus, because the court as a whole is perceived as legitimate, political actors will comply with its decisions even when they do not like them.

How do courts build and maintain legitimacy? In a study of the high courts of eighteen European Union member states, Gibson et al. find that building institutional legitimacy requires the creation of specific support among successive, non-overlapping constituencies.\textsuperscript{25} In order to be able to issue an unpopular decision, a court needs to render a range of popular decisions first.\textsuperscript{26} Scholars of American politics have observed that this insight also explains the behavior of the U.S. Supreme Court, which only rarely issues decisions that are truly counter-majoritarian in nature. Comparative scholars have reached similar conclusions about constitutional courts in Russia, and Germany, among others.\textsuperscript{27}

\subsection*{2.2.1 Backlash}

While both of these explanations—coordination benefits and legitimacy—suggest that judges can be important enforcers of human rights, they also suggest that judicial power is not unlimited. A key problem for judges is the possibility of backlash—a negative reaction to a judicial decision, formal or informal, that can limit a court’s ability to perform its obligation in the future.\textsuperscript{28} The potential for backlash is a constraint on the ability of the courts to effectively enforce the country’s human rights commitments. Courts therefore act strategically to reduce the risk of such backlash.

\begin{itemize}
\item \textsuperscript{19} Hardin 2003, Hardin 2013, 52; Hardin 1989, 101.
\item \textsuperscript{20} Hardin 2013, 57; Levinson 2011, 708.
\item \textsuperscript{21} Easton 1965, 1975; Gibson et al. 1998.
\item \textsuperscript{22} Easton 1965, 273; 1975, 444.
\item \textsuperscript{23} Easton 1975, 437.
\item \textsuperscript{24} Gibson & Caldeira 1995; Gibson et al. 1998; Lupu 2013a, 440-45.
\item \textsuperscript{25} Gibson et al. 1998, 355.
\item \textsuperscript{26} Gibson et al. 1998, 355; Shapiro 1998, 11.
\item \textsuperscript{27} Epstein et al. 2001; Vanberg 2005.
\item \textsuperscript{28} Fontana and Braman 2014; Alter 2000; Helfer 2002.
\end{itemize}
As Alter et al. observe in the introduction to this special issue, the purpose of “backlash politics” is to “overturn a precedent, eliminate or circumvent a legal ruling, or avert future losses in similar cases.”

Backlash can take the form of criticism of the court’s decision by the public, short- or long-term mobilization by civil society groups to offset or counteract a judicial decision, or action by another branch of government. For example, when the Judicial Committee of the Privy Council issued a series of unpopular opinions that effectively abolished the death penalty in three Caribbean countries that had retained the Privy Council, these countries withdrew from their human rights treaty commitments entirely.

When backlash comes in the form of formal action by other governmental actors directed at the court, it is sometimes referred to as court curbing. For example, when the U.S. Supreme Court repeatedly struck down social welfare legislation during the Lochner era, President Roosevelt responded with a court-packing plan that would alter the composition of the court and bring it in line with popular majorities. In a remarkably similar episode in India, the Indian parliament responded to counter-majoritarian social welfare decisions by adopting numerous constitutional amendments that overruled the courts’ constitutional interpretations.

Where courts face backlash, they can lose their legitimacy and thus their ability to compel compliance with their rulings going forward. Recent controversies involving decisions by constitutional courts in Poland, Hungary, Romania, Pakistan, Sri Lanka, South Africa, and Israel show the potential for backlash and how they hurt the ability of courts to enforce rights. In some cases, a court may have built up sufficient legitimacy over time to withstand backlash, but this is not always the case. As Tom Clark (2009, 974) argues, “political attacks on the Court serve as signals of a lack of specific support for the Court, which in turn indicates that further judicial recalcitrance will not be tolerated and that the Court will not be able to effectively set policy.” Backlash might also mean that the strong negative response to the judicial ruling outweighs the coordination benefits that courts bring.

As Alter et al. note, backlash occurs in the fourth and final phase of judicialized politics, and is part of the “strategies and actions that follow from a legal victory or loss.” In addition, the possibility of backlash affects earlier phases of judicialized politics, including the adjudication and compliance phases. Because the potential for backlash imposes a constraint on courts, they often act strategically to avoid it. Both public opinion and appeals from interest groups have been shown to affect judicial decision making. For example, the U.S. Supreme Court is more likely to strike down federal laws when public support for the court is greater. Judges also

29 Alter, Hafner-Burton, and Helfer 2016.
30 Helfer 2002.
31 This episode famously ended with Justice Owen Roberts’ switch in time in response to Roosevelt’s court-packing plan. Cushman 1994, 208-09.
32 Van der Walt 1999; Mate 2013.
33 Gardbaum 2015.
34 Alter, Hafner-Burton, and Helfer 2016.
35 Vanberg 2005; Carrubba 2009.
37 Clark 2009.
choose the precedents they cite based, in part, on the preferences of other actors both within and outside the court.\textsuperscript{38} They strategically attempt to set precedents that they believe will be followed and implemented by other actors, including political actors.\textsuperscript{39}

Courts also use legal reasoning to avoid backlash. Specifically, creating the perception that courts are apolitical is crucial to avoiding backlash. When individuals believe judges make decisions based on political, rather than legal, reasons, they are less likely to abide by those decisions.\textsuperscript{40} By contrast, when individuals believe judges make decisions based on legal doctrine, they are more likely to support those decisions.\textsuperscript{41} Indeed, among the key “checklist” items for international judicial effectiveness described by Laurence Helfer and Anne-Marie Slaughter (1997) are awareness of the court’s audience, incrementalism, and maintaining an appearance of impartiality—all of which are intended to minimize the potential for backlash.\textsuperscript{42} The same has been noted for domestic courts. As Martin Shapiro (1981, 34) noted, “judicial lawmaking and judicial independence are fundamentally incompatible,” meaning that political actors will not let judges get away with appearing to make law. An important goal for courts, therefore, is to avoid the appearance of law-making or “legislating from the bench”.

As we will show in the next section, one way to avoid the appearance of law-making is by using legal mechanism that avoid direct confrontation with the political branches and are more incremental in nature. Specifically, we will show that where judges do not invalidate domestic laws directly, but interpret them in conformity with a treaty, their decisions might be seen as more legitimate.

3. Theory: Enforcing International Human Rights Treaties in Domestic Courts

In this section, we develop our theory regarding domestic judicial enforcement of international human rights commitments. Our argument highlights both the relationship between the various phases of judicialized politics and the relationship between domestic institutions and international judicialization. We begin by describing two legal mechanisms by which national courts may apply international human rights treaty commitments: the direct effect mechanism and the canon of interpretation mechanism. We also explain how national legal doctrines determine the availability of each mechanism to domestic courts. We argue that while the direct effect mechanism faces significant legal obstacles, courts can often use the seemingly weaker canon of interpretation mechanism to great effect. We then explain how, in addition to avoiding these legal obstacles, the canon of interpretation mechanism also minimizes the risk of backlash against the courts’ decisions, making it a more politically viable tool.

3.1 Legal Mechanisms and Legal Constraints

In 2009, Madonna Louise Ciccone—better known as the singer Madonna—applied to adopt Chifundo James, an infant living in rural Malawi. But she faced a seemingly insurmountable

\textsuperscript{38} Lupu and Voeten 2012.
\textsuperscript{39} Knight and Epstein 1996.
\textsuperscript{40} Staton 2010.
\textsuperscript{41} Gibson, Caldeira & Spence 2005; Danner 2003.
\textsuperscript{42} Alter 2000.
legal obstacle. The country’s law prohibited adoption by “any applicant who is not resident in Malawi.” The court denied the application. When the case reached the Malawi Supreme Court of Appeal, Madonna’s lawyers invoked the U.N. Convention on the Rights of the Child and the African Charter on the Rights and Welfare of the Child. These human rights treaties, both ratified by Malawi, required that adoption decisions be made in the best interests of children, including through international adoption where appropriate. But again, a major legal obstacle stood in the way. Under Malawi’s constitution, treaties are not part of domestic law unless and until Parliament adopts implementing legislation.

Because the treaties had not been implemented, Malawi’s courts could not apply them directly to override the adoption law. Madonna’s lawyers advanced an alternative argument: the court should apply a legal doctrine known as canon of interpretation, under which domestic law must be interpreted in conformity with the country’s treaties. The court agreed: it stated that “the courts will try as much as possible to avoid a clash between what our laws say … and what the international agreements or conventions are saying.” It held that the term “resident” ought not to be given its ordinary legal meaning, but could include someone who had “a targeted long term presence aimed at ameliorating the lives of more disadvantaged children in Malawi.” Thus the court, even though it lacked the constitutional power to enforce human rights treaties directly and displace domestic law, reached the same result by reshaping the meaning of the adoption law to comply with the treaties. It allowed the adoption to proceed.

The Madonna adoption case illustrates how domestic courts can use different legal tools to give effect to human rights treaties. When a state ratifies a treaty, it is bound under international law to comply with its provisions. Yet, this does not automatically mean that the treaty becomes part of the country’s domestic law and is enforceable by its courts. The effect of treaties in the domestic legal system is governed by constitutional and public law doctrines that vary across countries. There are two principal mechanisms under which domestic courts may be empowered to apply treaties: they may give them direct effect, or use canons of interpretation to interpret domestic law in conformity with treaties.

The first mechanism, direct effect, means that “domestic law authorizes a domestic court to apply that treaty provision as a rule of decision.” In other words, the court may sanction a violation of the treaty much like a violation of a statute or other domestic law. The plaintiff brings a case alleging a treaty violation; the court hears the evidence and arguments; if it finds a violation, it grants a remedy. The remedies may include monetary damages, issuing an injunction.

---

43 Adoption of Children Act Chapter 26:01, Re, Ciccone, Appeal judgment, MSCA Adoption Appeal No 29 of 2009, [2009] MWSC 1, ILDC 1345 (MW 2009), 12th June 2009, Supreme Court of Appeal

44 The distinction between these mechanisms is distinct from the concept of judicial independence, which focuses on the relationship between the judiciary and other branches of government. A fully independent judiciary may not have either of these mechanisms available to it, depending on domestic legal doctrine. Likewise, these mechanisms also differ from the classification of legal systems (e.g., common versus civil law) that are often discussed in the literature, although the direct effect mechanism is not available to courts in most common law countries.

45 Sloss 2009, 11.
ordering the violation to cease, or invalidating state action inconsistent with the treaty. If, in addition to being given direct effect, treaties are considered superior to domestic laws, the court may also invalidate statutes inconsistent with the treaty.

For example, in 1997 the Bulgarian Supreme Court of Cassation heard a challenge to an article of the Criminal Procedure Law that prohibited courts from releasing prisoners prior to trial in several circumstances, including recidivism. Bulgaria was a party to the ICCPR, whose Article 9(3) provides that “it shall not be the general rule that persons awaiting trial shall be detained in custody.” Under the Bulgarian Constitution, ratified treaties become part of domestic law and “shall have primacy over any conflicting provision of the domestic legislation.” Because the Criminal Procedure Law categorically mandated pre-trial detention, the court held that it contradicted Article 9(3) and was preempted.

The second mechanism consists of applying a canon of interpretation under which domestic law must be interpreted to comply with the state’s treaty obligations. In the United States, this is often called the “Charming Betsy” doctrine, after a 19th-century case in which the Supreme Court held that “an Act of Congress ought never to be construed to violate the law of nations if any other possible construction remains.” Instead of applying the treaty itself as a rule of decision, the court applies domestic law as interpreted in light of the treaty. Under this mechanism, the treaty’s effect is derivative rather than direct: it cannot apply in the absence of some domestic law on point, and it cannot displace domestic law that clearly contradicts it. Thus, the canon of interpretation is in principle a legally weaker mechanism, giving the court less legal power to override treaty-inconsistent state action.

Several additional observations regarding these mechanisms are worth noting because they inform our argument about their political effects, which we develop in Section 3.2 below. First, legal rules that govern the status of treaties in domestic law may make either or both mechanisms unavailable. The direct effect mechanism is unavailable in countries which, like Malawi, follow a “dualist” approach under which ratified treaties are not part of domestic law and cannot be enforced by domestic courts without implementing legislation. This approach is common: as of 2012, treaties can only have direct effect in 55% of the countries in our sample. By contrast, other countries such as Bulgaria follow a “monist” approach, under which ratified

---

46 Prosecutor General, Ruling on a proposal to adopt an interpretative decision, Case No 1/1997, Ruling No 1, ILDC 611 (BG 1997).
47 Bulgarian Constitution, Art. 5(4).
48 The Court then deferred the case to the Constitutional Court, which had jurisdiction to formally declare the unconstitutionality of the relevant provision. Before that case was heard, the Criminal Procedure Law was reformed. In a later case, the Court used similar reasoning to conclude that the state could be held liable for violating the ICCPR right to be tried without undue delay. Encheva v Pazardzhik District Investigation Service, Cassation appeal, Judgment no 719, Civil Case no 2397/2001, ILDC 613 (BG 2002).
49 Murray v. The Schooner Charming Betsy, 6 U.S. (2 Cranch) 64, 118 (1804).
50 For example, the highest British court has held that the executive’s power to ratify treaties “does not extend to altering the law or conferring rights on individuals … without the intervention of Parliament.” J.H. Rayner (Mincing Lane) Ltd. v. Department of Trade and Industry, [1990] 2 A.C. 418, 500.
51 Verdier & Versteeg 2015.
treaties become part of domestic law without implementing legislation. In such countries, courts are in principle empowered to give direct effect to human rights treaties.

Likewise, the canon of interpretation mechanism is recognized in some countries but not others. In several countries, including the United States, well-established judicial precedent requires courts to apply the canon to avoid conflicts between treaties and domestic law. In some countries, the canon is codified in the constitution itself. Article 233 of the South African constitution provides that “[w]hen interpreting any legislation, every court must prefer any reasonable interpretation of the legislation that is consistent with international law over any alternative interpretation that is inconsistent with international law.” According to our data, the canon of interpretation is recognized in 65% of countries as of 2012. Many of these countries are dualist, which means that even though courts cannot give human rights treaties direct effect, the canon of interpretation may provide an alternative mechanism through which courts can use those treaties to shape the interpretation of domestic law.

Second, courts retain substantial discretion in deciding whether to invoke each mechanism even when it is legally available. This point may seem surprising, because on their face the doctrines often speak imperatively. Under the Bulgarian constitution, treaties “shall have primacy” over domestic law, which seems to indicate that courts are bound to apply them. Nevertheless, courts have developed numerous doctrines they can invoke to avoid giving direct effect to treaties. For example, courts in virtually all monist systems recognize a distinction between “self-executing” and “non-self-executing” treaties. Courts in monist countries often refuse to give direct effect to human rights treaties on that ground. For example, despite a constitutional provision making treaties part of Austrian law, the Austrian Supreme Court held in 2008 that the ICCPR was non-self-executing and refused to give it direct effect. Beyond the self-execution doctrine, domestic courts can also invoke several other “avoidance doctrines” to decline to give direct effect to a treaty.

The canon of interpretation, as articulated in most countries, also appears to speak imperatively. In South Africa, courts “must prefer any reasonable interpretation of the legislation that is consistent with international law.” Yet, this instruction only applies when such a “reasonable” interpretation is available. If the interpretation consistent with international law stretches the legislation’s meaning beyond reason, the canon does not apply and the domestic legislation prevails. Again, because reasonableness is an imprecise standard, the court retains substantial discretion in deciding whether to apply the canon. In the Madonna case, the Zambian court went quite far in reinterpreting the term “resident” to fulfill the treaty’s command. Other courts have been less bold. In sum, even when one or both of the mechanisms are available to a court, it may choose to invoke either or none.

---

52 Nollkaemper 2011; Shelton 2011; Verdier & Versteeg 2015.
53 Nollkaemper 2011.
54 Perterer v Land Salzburg and Austria, Appeal judgment, 1Ob8/08w, ILDC 1592 (AT 2008), 6th May 2008, Supreme Court of Justice [OGH]
55 Benvenisti 1993.
Third, although the canon of interpretation is a legally weaker tool than direct effect, it can nonetheless be effective in fulfilling treaty obligations. Courts can deploy the canon in multiple ways to reshape domestic law and enforce treaty rights, providing “a powerful alternative” to direct effect. The following examples illustrate such applications.

The most straightforward application is for courts to apply the canon to privilege an interpretation of domestic law that conforms to human rights treaties. In some cases, this allows courts to punish clear rights violations. After the reunification of Germany, East German border guards were prosecuted for killing a man attempting to cross the Berlin Wall. The guards argued that their actions were permitted by East German laws and by their orders, under which it was “better the fugitive dies than the escape succeeds.” The Federal Supreme Court held that, because the GDR ratified the ICCPR, the canon required that its laws be interpreted in light of the treaty rights to life and to leave one’s country, even though the ICCPR did not have direct effect. As a result, the court rejected the border guards’ justification for the killing. Beyond punishing or deterring violations, domestic courts can use the canon to reform domestic law in light of human rights treaties, as in the Madonna case.

Domestic courts reviewing the legality of administrative decisions can also invoke the canon to use human rights treaties as a benchmark. In reviewing an administrative decision to deport a father convicted of drug trafficking, the Australian High Court held that because Australia had ratified the Convention on the Rights of the Child, immigration authorities had to consider the “best interests of the child.” Courts in Canada, New Zealand, and Hong Kong have also relied on human rights treaties in reviewing administrative action. Importantly, common law countries, where treaties usually lack direct effect, often have a strong tradition of judicial review of administrative action under which this approach can be deployed.

Finally, domestic courts can deploy the canon to interpret the constitution itself. Used in this way, the canon can be a powerful tool in countries with a constitutional bill of rights, effectively empowering courts to displace legislation inconsistent with human rights treaties even when they lack direct effect. In 2007, Canadian unions challenged a provincial law restricting collective bargaining in health and social services. The challenge was based on s. 2(d) of the Canadian Charter of Rights and Freedoms, which protects freedom of association. The Supreme Court interpreted s. 2(d) to protect collective bargaining against undue interference and held the law unconstitutional. It based this interpretation on several human treaties ratified by Canada but which lacked direct effect. It reiterated that “the Charter should be presumed to provide at least as great a level of protection as is found in the international human rights documents that Canada has ratified.”

---

56 Waters 2006; Nollkaemper 2011, 139, 143-44.
58 Courts in common law jurisdictions can also use human rights treaties as a guide to reform common law rules. The New Zealand Supreme Court has drawn on the ICCPR to develop new common law rules governing breach of privacy. Waters 2006, 670.
The Supreme Court of Bangladesh adopted a similar approach in a case alleging that an education official had insulted and threatened a headmistress for failing to wear a headscarf. The court held that attempts to coerce or impose a dress code on women without legal authority amounted to sexual harassment, which was a discriminatory act violating the constitution and international standards. It also held that “arbitrary and abusive gender-based codes” violated the rights to privacy and self-expression protected by human rights treaties. As a dualist country, Bangladesh did not give these treaties direct effect. However, the court stated that it could “look into these conventions and covenants as an aid to interpretation of ... [the constitutional bill of rights].”

The Court ordered the Ministry of Education to ensure that women working in schools are not subject to harassment or required to wear hear covering.

3.2 Political Effects of Legal Mechanisms

In this section, we develop our argument about the relative political effectiveness of the legal mechanisms discussed above. Is one of the mechanisms a more politically effective tool for enforcing human rights treaties? To answer this question, we build on three insights discussed above. First, although the canon of interpretation mechanism is a legally weaker form of review, courts can use it to political effect. Second, although judges can choose which mechanism(s) are legally available to them, they can choose whether or not to employ those mechanisms. Third, judges seek to avoid political backlash when possible. Courts are concerned with maintaining their legitimacy and their ability to compel compliance with their decisions more broadly. Backlash to judicial decisions threatens the legitimacy of domestic courts and their ability to enforce international human rights law

Although legal mechanisms for enforcing human rights treaties are often available to national courts, judges will use their discretion to choose which mechanism, if any, to employ, based in part on whether or not doing so may result in political backlash. If only one mechanism is legally available, judges may decline to issue a treaty-enforcing ruling if they risk of backlash for using that mechanism is too great. If both mechanisms are available, judges may choose to use the one they expect will be less likely to result in backlash.

The relative effectiveness of these mechanism therefore turns, part, on which one is more likely to result in backlash. In order to avoid backlash, courts can resort to various techniques that de-politicize their decisions and show self-restraint. When courts strike down domestic laws based on the constitution or a treaty, their action is likely to be perceived as an arbitrary exercise of judicial power, especially if the constitution or treaty is ambiguous. At the most general level, courts will attempt to avoid creating the perception that they are usurping the legislature’s law-making function. Since some rights provisions of the ICCPR and other human rights agreements are notoriously ambiguous, and lack precision, interpreting them and relying on such interpretations to strike down domestic laws makes courts especially vulnerable to this charge.

---

62 It is for this reason that there is a growing trend among constitution-makers to draft rights in greater detail. See Versteeg & Zackin forthcoming.
With respect to human rights treaties, we argue that use of the direct effect mechanism is more likely to be viewed as an abuse of judicial power and thus is more likely to cause backlash than the use of the canon of interpretation mechanism. Unlike the direct effect mechanism, by which domestic courts openly enforce human rights treaties against the political branches by striking down laws, invalidating government actions, or providing other remedies, the canon of interpretation “is essentially a technique to shift power in a subtle way to the courts.”

Even when direct effect is legally available, by using the canon domestic courts can also avoid bypassing or overriding domestic laws. As Nollkaemper (2011, 145) observes, “it is primarily a rule of national law that is applied and the courts accordingly do not usurp any legislative powers.”

There are various reasons why such decisions reduce the likelihood of backlash. First, the use of the canon of interpretation is less likely to be perceived as a political decision. Striking down a domestic law based on the ICCPR produces a direct confrontation with the political branches, and might be seen as judges venturing into the political realm, especially when ICCPR rights lack precision and leave substantial room for discretion. Even when domestic courts are empowered to directly apply international human rights treaties to displace domestic law, doing so directly challenges the lawmaking authority of the political branches, potentially inviting backlash. In effect, it is similar to exercising constitutional judicial review, the strong exercise of which has produced backlash in many countries.

By contrast, using a canon of interpretation relies on a core judicial function: declaring what the law means. It is therefore more likely to be perceived as legitimate by politicians and the public. The canon of interpretation mechanism can avoid direct confrontations with other actors, while still allowing the court to effect justice in the individual case.

Comparative scholars have observed that courts commonly resort to various avoidance canons, through which they can read legislation so that it complies with the constitution rather than invalidating it directly. Alexander Bickel advocated that courts exercise “passive virtues” by avoiding deciding cases on broad constitutional grounds when narrower grounds can be found. For example, in a famous decision, the Indian Supreme Court did not invalidate the highly contentious “Muslim Women’s (Protection of Rights to Divorce) Act”, but instead interpreted it in conformity with the Indian Constitution, thereby avoiding some of the strong political backlash that had surrounded its earlier decisions that balanced gender equality and religion. Other scholars have observed that courts can rely on deferral techniques, by which judicial decisions striking down laws are not given immediate application in order to allow the political branches to reform them. Stephen Gardbaum has suggested that constitutional courts

---

64 Nollkaemper 2011, 145.
66 Delaney 2016.
68 Hirschl 2013, 434. As another example, the Turkish Constitutional Court attempted to avoid backlash by reading a highly-contested law on headscarves in higher education in conformity with the constitution rather than striking it down as unconstitutional. Roznai and Yolcu 2012, 180-81.
69 Dixon & Issacharoff 2015.
in young democracies might prefer a “weak form” of judicial review in which legislatures can override judicial rulings, which produces dialogue between the political and judicial branches. Mark Tushnet has similarly proposed weaker forms of judicial review to enforce social and economic rights. These techniques are means for courts to advance their objectives while avoiding direct clashes with the political branches that might provoke backlash and compromise their independence.

Second, the strongest remedy typically available under direct application is to declare a domestic law void for inconsistency with a human rights treaty. Such a declaration affects all future cases involving that law, potentially including cases having nothing to do with the human rights violation. It puts the onus on the political branches to react, hopefully by reforming the law consistently with international human rights—but they can also reaffirm the law, setting the stage for a political confrontation with the court. By contrast, the canon of interpretation tends to be used on a case-by-case basis: the law is not invalidated, but is read so as to avoid conflicts with the treaty in cases where such a conflict would arise. This does not necessarily require a response by the legislature or government, thus reducing the risk of direct confrontation. By re-interpreting domestic law on a case-by-case basis, courts can thus take an incremental approach to reforming domestic law in line with human rights treaties while avoiding direct confrontations with the political branches.

Third, even if other actors object to the court’s interpretation, the use by courts of the canon of interpretation is difficult to circumvent. As far as we are aware, no legislature has generally prohibited courts from taking ratified treaties into account in interpreting domestic law, and such a prohibition would be difficult to enforce. The legislature can amend laws on a case-by-case basis to reverse interpretations arrived at by courts using the canon, but this is a labor-intensive approach, and since the law itself is still in force, the political impetus to address the issue will likely be weaker. In this sense, consistent interpretation exploits the passivity of legislatures not faced with an emergency.

These arguments lead to our key hypothesis regarding human rights treaties. The larger probability of backlash associated with direct treaty application suggests that, even when this mechanism is available, courts may avoid the case or issue a ruling that does not enforce the treaty. If the direct effect mechanism is available but the canon of interpretation is not, the court’s fear of backlash should diminish the extent to which it relies on direct effect and thus the extent to which it enforces such treaties. By contrast, if the canon of interpretation is available, the court will be more likely to be able to issue a treaty-enforcing ruling without fear of backlash.

Thus, while direct effect may seem the more powerful legal mechanism in the adjudication phase, the potential political constraints on that mechanism in the feedback phase

---

70 According to Gardbaum, these weaker forms of review guard the independence of the court in the long-run. Gardbaum 2015, 289. See also Gardbaum 2013, 21-46.

71 Tushnet 2008.

72 Some U.S. states have proposed legislation banning the use of Shariah law or international law by their courts; it is unclear whether this would prohibit the use of ratified treaties to interpret domestic law.
result in the canon of interpretation mechanism more effective in applying human rights treaties. This leads to our conceptual hypothesis, which we operationalize in Section 4 below:

Conceptual Hypothesis: The canon of interpretation is more effective than direct effect as a mechanism of human rights treaty enforcement.

By contrast, our arguments do not lead to specific predictions regarding the extent to which we should or should not observe judicial backlash. Strategic courts attempt to minimize backlash, so the extent to which we observe backlash is a product of multiple factors, including such strategies. Indeed, it might be the case that instances of backlash in response to the direct application of treaties are uncommon, but that could be precisely because courts avoid using that mechanism when the potential for backlash is greater. Likewise, there is no guarantee that using a canon of interpretation prevents backlash entirely. For example, if the canon were used in such a way that ignores the clear meaning of a domestic statute’s text or in a way that stretches the meaning of a treaty provision, a court may still be seen as engaged in law-making. Also, courts cannot perfectly anticipate their audiences’ reactions. In some cases, courts may underestimate the possibility of backlash to their decisions or may take a calculated risk that does not pay off. Finally, in some instances, courts may be willing to take the risk of backlash because the legal or normative stakes in a case are sufficient large and the court believes its legitimacy can weather the storm of a backlash in the long run, possibly because the court has sufficient diffuse support. Nevertheless, if the court can enforce the law while minimizing the likelihood of backlash, it will often seek to do so.

4. Research Design

4.1 Key Variables and Operational Hypotheses

4.1.1 ICCPR

We focus on the ICCPR because previous studies have suggested that courts are important in the enforcement of the ICCPR. In addition, we can draw on an extensive literature on human rights treaty effectiveness, which has made substantial progress in developing methods to identify the impact of ICCPR ratification. Moreover, the ICCPR includes a range of different rights commitments, which allows us to test the treaty’s impact on a range of different variables. The ICCPR was adopted in 1966, and entered into force in 1976. As of 2016, 168 states are parties to the ICCPR.

4.1.2 Measures of Canon of Interpretation and Direct Effect Mechanisms

To measure whether the canon of interpretation and direct effect mechanisms are available to domestic courts, we coded an original data set. Existing initiatives to map international law in domestic legal systems are less systematic and/or more limited in scope. Specifically, Hathaway (2008) and the Comparative Constitutions Project assemble information

73 Lupu 2013a; Simmons 2009; Powell and Staton 2009.
74 Simmons 2009; Hill 2010; Lupu 2013b.
75 Ginsburg et al. 2008.
on the domestic status of international law based on the constitution alone. Yet, many relevant aspects of a domestic legal system’s relationship to international law are not found in the text of the constitution but in ordinary legislation, case law, practice, and executive orders.76 Our data set is based on coding of all these different legal sources.

In collecting our data, a first step was to identify and define the substantive issues that define a state’s relationship with international law. We identified about 50 issues, including a range of issues that allow domestic courts to use treaties in domestic cases. For each country, we commissioned a written memorandum that provides a narrative answer to each of the questions, and documenting how this answer has changed over time. These memorandums prepared by the principal investigators as well as a number of research assistants and consultants, primarily foreign students with prior knowledge of the relevant legal systems. The memorandums not only document the answer to each question, but also the sources in answering the question, thus indicating whether the basis for a given answer is found in the constitution, ordinary legislation, court decisions, executive interpretations or secondary sources. 77 For most countries, the constitution was the starting point for the memorandum, yet it was never the sole basis for the answers. While the authors of the memos were instructed to settle upon one ultimate answer to each question, they were also asked to indicate if there existed multiple interpretations, and to elaborate on those. Where multiple interpretations existed, we relied on the most authoritative source of that system to make a judgment call. All judgment calls were made by the principal investigators.

To create quantitative data, we hand-coded these memoranda. All coding was done by the authors. From this data set, we draw information on how domestic courts have dealt with international law. We use data on the following questions:

1) Do ratified treaties automatically become part of domestic law without implementing legislation?
2) If the answer to (1) is yes, what is the relationship of treaties to ordinary statutes (superior, equal, inferior)?
3) Does the constitution expressly refer to international human rights treaties (e.g., the ICCPR or the European Convention on Human Rights)? If so, which one(s) and are they formally incorporated into domestic law?
4) Is there a rule or presumption that domestic statutes (and/or other law) should be interpreted in conformity with obligations under ratified treaties?

From these questions, we constructed two binary indicators. The first captures the degree to which international treaties are directly available in domestic courts and can be used to strike down domestic legislation (DIRECT EFFECT). Specifically, we combined answers 1 through 3 into a single binary indicator DIRECT EFFECT that takes the value 1 if countries apply treaties directly and:

---

76 In addition to quantitative coding of constitutions, several qualitative surveys offer more in-depth information. However, such surveys cover a limited number of countries (Shelton 2011) or specific issues (Council of Europe 2001; Hollis et al. 2005; Sloss 2009), and do not reflect change over time.

77 We thank the Comparative Constitutions Project for providing us access to their historical repository of constitutions.
designate them as having higher status than domestic legislation or if they have incorporated an international human rights agreement directly into their constitution.

The second variable we coded captures the degree to which courts can interpret domestic legislation in line with ratified treaties. This indicator CANON OF INTERPRETATION is coded as 1 if the relevant legal system recognizes such a canon of interpretation. It is possible for both variables to be coded as either 1 or 0 in a given country-year.

Figure 4 describes trends in the availability of these mechanisms to domestic courts over time. The percentage of states in which treaties have direct effect has increased substantially over the past two centuries. By contrast, the total number of countries that use that canon has increased over time, although the percentage of such states has remained fairly constant. Part I of the Supplementary Information describes the countries in which the two mechanisms are available.

Figure 4: Prevalence of Direct Effect and Canon of Interpretation Mechanisms over Time
4.1.3 Dependent Variables

As dependent variables, we follow Lupu (2013a) by adopting the measures of Freedom of Assembly and Association, Freedom of Speech, and Freedom of Religion provided by the Cingranelli-Richards Human Rights Data Project. Each of these measures is coded as 0, 1, or 2 for each country-year. A score of 2 indicates the applicable freedom was not restricted in that year, while a score of 0 indicates it was severely restricted. We test our hypotheses using a series of ordered probit models.

4.1.4 Operational Hypotheses

In Section 3.2, we stated that our general hypothesis is that the canon of interpretation is more effective than direct effect as a mechanism of human rights treaty enforcement. Operationalizing this in light of the key measures we are using our empirical analysis yields the following operational hypotheses:

H1: The extent to which the canon of interpretation increases the effect of ICCPR ratification on freedom of association is greater than the extent to which the direct effect mechanism increases the effect of ICCPR ratification on freedom of association.

H2: The extent to which the canon of interpretation increases the effect of ICCPR ratification on religious freedom is greater than the extent to which the direct effect mechanism increases the effect of ICCPR ratification on religious freedom.

H3: The extent to which the canon of interpretation increases the effect of ICCPR ratification on free speech rights is greater than the extent to which the direct effect mechanism increases the effect of ICCPR ratification on free speech rights.

4.2 Hypothesis Testing

Estimating the effects of treaty commitments is known to be difficult. Governments select the treaties they join in part based on their interests and the extent to which they expect to conform their behavior to the treaties' requirements. Preferences drive a significant part of the difference between treaty-members and non-members: members may be significantly more likely to join a treaty simply because they prefer the policy choices embodied in it.78 States that join a treaty may not be comparable to states that refrain from doing so. No methodology, quantitative or qualitative, can allow us to infer causation based on these observational data without assumptions. We design our analysis to weaken these assumptions as much as possible.

Following much of the recent literature, we use a propensity score matching approach. Propensity score matching reduces the differences between the group of states that ratified the treaty and the group of states that did not. It does so by pairing observations that are as similar in every relevant way as possible except that one has ratified the treaty while the other has not.

---

78 Downs et al. 1996.
If the observations are similar along all relevant dimensions except that one has ratified the treaty, then observed differences in the dependent variable can be attributed to the treaty. This approach estimates each state’s probability of treaty ratification based on factors that predict ratification, and then matches treaty members to treaty nonmembers based on this probability. Among the advantages of this approach are that it creates covariate balance and weakens distributional assumptions.

A notable limitation of matching, however, is that it relies on conditioning exclusively on observable variables. If an unobservable (or unmeasured) factor affects treaty commitment decisions and is not included in the matching model, this can bias inferences in a manner analogous to omitted-variable bias in a standard regression context. In the context of human rights treaty ratification, Lupu (2013b) argues that a key latent (or difficult to observe) factor is states’ preferences for treaty commitments. As Downs et al. (1996) and others have noted, without controlling for underlying preferences, we cannot distinguish whether compliance with international institutions results from these preferences or whether commitment to an institution affects the probability of compliance. To address this issue, we include in our models the measure of Treaty Commitment Preferences developed by Lupu (2013b), which is an annual estimate of each country’s probability of ratifying the ICCPR. Detailed information on how this measure is constructed can be found in Part II of the Supplementary Information.

We match ICCPR members to non-members based on variables that predict ICCPR commitments, including Treaty Commitment Preferences and several other variables. As a measure of regime type, we use the Polity IV data (POLITY). We include a measure of judicial independence provided by the CIRI project (JUDICIAL INDEPENDENCE) which is coded 0 for “not independent,” 1 for “partially independent,” and 2 for “generally independent.” Newer regimes and well-established regimes may have different preferences with respect to treaty commitment and different incentives with respect to human rights abuses. We control for this factor using the Polity IV data (REGIME DURABILITY). As a measure of interstate wars, we use data from the Correlates of War project (INTERNATIONAL WAR). For civil conflicts, we use data from the UCDP/PRIO armed conflict database (CIVIL WAR). We control for economic development using a measure of the natural log of per capita GDP provided by the World Bank (GDP PER CAPITA).

To address potential differences among states of different sizes and potential monitoring biases based on this factor, we include the natural log of a state’s population, using data provided by the World Bank. We also include a measure of the number of international NGOs (INGOS) in a country using the data provided by Hafner-Burton and Tsutsui (2005). We include the year of the observation in the matching model because treaty commitments may have been more likely in some years than others. Finally, we include in the matching model the Direct Effect and Canon of Interpretation variables discussed in the next subsection. There are many units with missing data among these variables. Because the underlying reasons for missing data are likely

79 These variables are consistent with recent similar analyses, e.g., Hill 2010; Lupu 2013a, 2013b, 2015; Chilton and Versteeg 2015a, 2015b.
nonrandom, listwise deletion of these observations may result in biased inference.\textsuperscript{81} We therefore impute the missing values using the Amelia II Program.\textsuperscript{82} Part III of the Supplementary Information provides additional information about this procedure.

ICCPR RATIFICATION is the dependent variable in the matching model. It is coded “1” if a state has ratified the ICCPR as of a given year and “0” otherwise. We match ICCPR members to non-members using the nearest-neighbor algorithm provided by the Matchit package in the R programming language.\textsuperscript{83} A country-year that has ratified the ICCPR is matched to another country-year that has not ratified the ICCPR if the two are estimated to have nearly the same probability of entering the treaty. This creates a matched pair that differs with respect to their ICCPR commitment, but differs very little in terms of their probability of joining the ICCPR.

We create a matched sample for each of our dependent variables. The matched samples differ slightly from each other because of differing availability of the dependent variable data. Parts IV and V of the Supplementary Information provides summary statistics and balance statistics for the matched samples. Figures 1-3 show the countries included in the matched samples. The darker the shade of gray in which a country appears, the greater the number of years in which it appears in the matched sample.

Because our hypotheses are about the conditions under which ICCPR RATIFICATION affects human rights practices, we test our hypotheses by interacting ICCPR RATIFICATION with CANON OF INTERPRETATION and DIRECT EFFECT. To address serial correlation, we include a lag of the applicable dependent variable for year \( t - 1 \). A Lagrange multiplier test indicates that additional lags are not necessary to address serial correlation. The temporal scope of our data includes the years 1982-2007. The CIRI data begin in 1981, but to include the lagged dependent variable we begin our analysis in 1982.

\textsuperscript{81} Little and Rubin 1987.
\textsuperscript{82} Honaker, King, and Blackwell 2011.
\textsuperscript{83} Ho et al. 2011.
Figure 1: Countries Included in Matched Sample - Association

Figure 2: Countries Included in Matched Sample - Religion
Figure 3: Countries Included in Matched Sample - Speech
5. Results

Table 1 provides the results of our models. Given the coding of the dependent variables, a positive coefficient indicates that an independent variable is associated with greater respect for rights, while a negative coefficient indicates that an independent variable is associated with greater repression. Several results are evident in the table. First, the extent to which ICCPR ratification improves respect for all three rights increases when the canon of interpretation mechanism is available (i.e., the coefficient of CANON OF INTERPRETATION * ICCPR RATIFICATION is significant and positive in all 3 models). Second, the extent to which ICCPR ratification improves respect for free speech rights also increases when the direct effect mechanism is available. By contrast, the extent to which the direct effect mechanism affects whether ICCPR ratification improves freedom of association and religion is statistically indistinguishable from zero.

To test our hypotheses, we actually need to compare the coefficients of the interaction terms to each other, rather than to zero. If our hypotheses are correct, the coefficient of CANON OF INTERPRETATION * ICCPR RATIFICATION should be significantly larger than the coefficient of DIRECT EFFECT * ICCPR RATIFICATION. While this is the case with respect to the freedom of association and religious freedom (p < 0.05), it is not so with respect to the freedom of speech. In other words, we cannot reject the null hypothesis that both mechanisms are equally effective with respect to the ICCPR and the freedom of speech. These results are robust to ordinary least squares estimation, the results of which are reported in Part VI of the Supplementary Information.

We further test our hypotheses by estimating (1) the marginal effects of ICCPR ratification on civil rights, (2) the extent to which the availability of the direct effect and canon of interpretation moderate those effects, and finally (3) whether these moderation effects are statistically significantly different from each other. We present the results of these estimations in Figures 5-7.84

Figure 5 focuses on the freedom of assembly and association. We plot 4 marginal effects. Point estimate A is the change in the probability of a state severely restricting these rights (i.e., CIRI category 0) associated with ICCPR ratification when the canon of interpretation mechanism is available. Point estimate B is the change in the probability of a state having unrestricted rights (i.e., CIRI category 2) associated with ICCPR ratification when the canon of interpretation mechanism is available.85 Point estimate C is the change in the probability of a state severely restricting these rights (i.e., CIRI category 0) associated with ICCPR ratification when the direct effect mechanism is available. Point estimate D is the change in the probability of a state having

---

84 Our visual presentation of interaction effects differs from the more common presentations in the style of Brambor et al. (2006) because both of our independent variables are binary, and our dependent variable is trichotomous.

85 We do not analyze the marginal effects of ICCPR ratification on the probability of a state being in the middle CIRI category because the direction from which this change comes would be ambiguous (i.e., a 5% increase in this probability could reflect a worsening or an improvement in rights).
unrestricted rights (i.e., CIRI category 2) associated with ICCPR ratification when the direct effect mechanism is available. Because point estimates A and B are statistically distinguishable from zero, we can conclude that ICCPR ratification improves respect for freedom of association rights when the canon of interpretation mechanism is available. Because point estimates C and D are not statistically distinguishable from zero, we cannot conclude that ICCPR ratification improves respect for freedom of association rights when the direct effect mechanism is available. Most importantly, point estimate A is statistically distinguishable from point estimate C (p < 0.05), and point estimate B is statistically distinguishable from point estimate D (p < 0.05). This indicates that the extent to which the canon of interpretation increases the effects of ICCPR ratification on freedom of association rights is greater than the extent to which the direct effect mechanism does so, as predicted by Hypothesis 1.

Figure 6 provides similar estimates of marginal effects with respect to freedom of religion. As above, we can conclude that ICCPR ratification improves respect for religious freedom when the canon of interpretation mechanism is available, but cannot reach this conclusion with respect to the direct effect mechanism. While point estimate A is not statistically distinguishable from point estimate C (p < 0.05), point estimate B is statistically distinguishable from point estimate D (p < 0.05). Cumulatively, the extent to which ICCPR ratification improves religious freedom when the canon of interpretation mechanism is available is significantly (p < 0.05) greater than the extent to which it does so when the direct effect mechanism is available, as predicted by Hypothesis 2.

Finally, Figure 7 provides similar estimates with respect to the freedom of speech and the press. These results differ in important ways from those provided by Figures 5 and 6. All four point estimates are statistically distinguishable from zero, meaning that ICCPR ratification improves the respect for free speech rights when either mechanism is available. Point estimate A is not statistically distinguishable from point estimate C (p < 0.05), and point estimate B is not statistically distinguishable from point estimate D (p < 0.05).

The coefficients of other independent variables are consistent with existing findings. Democracies are more likely to respect civil rights. Violations of rights are more likely during civil wars, although this is only statistically significant with respect to freedom of speech. While civil wars are known to be strong predictors of human rights violations, analysts tend to analyze this relationship with respect to physical integrity rights. It may be the case that civil wars have more ambiguous or inconsistent effects on civil rights violations. Instances of rights violations are more likely, all else equal, in countries with larger populations, as indicated by the negative coefficients we estimate for these variables.

---

86 See, e.g., Hill and Jones 2014.
Table 1: Ordered Probit Models of Civil Rights

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Association</td>
<td>Religion</td>
<td>Speech</td>
</tr>
<tr>
<td>ICCPR Ratification</td>
<td>-0.336</td>
<td>-0.247</td>
<td>-0.397*</td>
</tr>
<tr>
<td></td>
<td>(0.236)</td>
<td>(0.241)</td>
<td>(0.227)</td>
</tr>
<tr>
<td>Canon of Interpretation</td>
<td>-0.0198</td>
<td>-0.183</td>
<td>-0.0233</td>
</tr>
<tr>
<td></td>
<td>(0.207)</td>
<td>(0.220)</td>
<td>(0.203)</td>
</tr>
<tr>
<td>Canon of Interpretation * ICCPR Ratification</td>
<td>0.760***</td>
<td>0.775***</td>
<td>0.511**</td>
</tr>
<tr>
<td></td>
<td>(0.257)</td>
<td>(0.268)</td>
<td>(0.240)</td>
</tr>
<tr>
<td>Direct Effect</td>
<td>0.213</td>
<td>-0.0451</td>
<td>-0.206</td>
</tr>
<tr>
<td></td>
<td>(0.189)</td>
<td>(0.208)</td>
<td>(0.171)</td>
</tr>
<tr>
<td>Direct Effect * ICCPR Ratification</td>
<td>-0.298</td>
<td>0.0863</td>
<td>0.483**</td>
</tr>
<tr>
<td></td>
<td>(0.265)</td>
<td>(0.273)</td>
<td>(0.244)</td>
</tr>
<tr>
<td>Polity</td>
<td>0.0750***</td>
<td>0.0415***</td>
<td>0.0601***</td>
</tr>
<tr>
<td></td>
<td>(0.0129)</td>
<td>(0.0115)</td>
<td>(0.0114)</td>
</tr>
<tr>
<td>Judicial Independence</td>
<td>-0.0121</td>
<td>0.300***</td>
<td>0.285***</td>
</tr>
<tr>
<td></td>
<td>(0.0950)</td>
<td>(0.0994)</td>
<td>(0.0986)</td>
</tr>
<tr>
<td>Regime Durability</td>
<td>-0.00188</td>
<td>-0.00108</td>
<td>0.00405**</td>
</tr>
<tr>
<td></td>
<td>(0.00192)</td>
<td>(0.00237)</td>
<td>(0.00187)</td>
</tr>
<tr>
<td>Civil War</td>
<td>-0.192</td>
<td>-0.208</td>
<td>-0.245*</td>
</tr>
<tr>
<td></td>
<td>(0.131)</td>
<td>(0.133)</td>
<td>(0.128)</td>
</tr>
<tr>
<td>International War</td>
<td>0.270</td>
<td>-0.452</td>
<td>-0.0603</td>
</tr>
<tr>
<td></td>
<td>(0.258)</td>
<td>(0.306)</td>
<td>(0.224)</td>
</tr>
<tr>
<td>GDP per Capita (logged)</td>
<td>0.00473</td>
<td>-0.194***</td>
<td>-0.00728</td>
</tr>
<tr>
<td></td>
<td>(0.0631)</td>
<td>(0.0578)</td>
<td>(0.0527)</td>
</tr>
<tr>
<td>Population (logged)</td>
<td>-0.0749</td>
<td>-0.208***</td>
<td>-0.0809*</td>
</tr>
<tr>
<td></td>
<td>(0.0479)</td>
<td>(0.0532)</td>
<td>(0.0436)</td>
</tr>
<tr>
<td>INGOs</td>
<td>7.29e-05</td>
<td>8.23e-05</td>
<td>0.000114</td>
</tr>
<tr>
<td></td>
<td>(0.000156)</td>
<td>(0.000160)</td>
<td>(0.000159)</td>
</tr>
<tr>
<td>Treaty Commitment Preferences</td>
<td>-0.0280</td>
<td>0.00284</td>
<td>0.00321</td>
</tr>
<tr>
<td></td>
<td>(0.227)</td>
<td>(0.234)</td>
<td>(0.215)</td>
</tr>
<tr>
<td>Year</td>
<td>-0.0587***</td>
<td>0.0218</td>
<td>0.00324</td>
</tr>
<tr>
<td></td>
<td>(0.0162)</td>
<td>(0.0153)</td>
<td>(0.0139)</td>
</tr>
<tr>
<td>Rights_{t-1}</td>
<td>1.697***</td>
<td>1.157***</td>
<td>1.199***</td>
</tr>
<tr>
<td></td>
<td>(0.112)</td>
<td>(0.0885)</td>
<td>(0.107)</td>
</tr>
<tr>
<td>Fixed Effects for Year</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>N</td>
<td>748</td>
<td>718</td>
<td>746</td>
</tr>
<tr>
<td>Log-Likelihood</td>
<td>-370.1</td>
<td>-443.1</td>
<td>-447.6</td>
</tr>
</tbody>
</table>

Robust standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1
Figure 5: Effects of ICCPR Ratification on Freedom of Association Rights
First Differences (95% Confidence Intervals)

Figure 6: Effects of ICCPR Ratification on Religious Freedom
First Differences (95% Confidence Intervals)
Figure 7: Effects of ICCPR Ratification on Free Speech Rights
First Differences (95% Confidence Intervals)
5.1. Discussion

Although we found support for H1 and H2, we cannot reject the null hypothesis with respect to H3. Both judicial mechanisms appear to be effective at enforcing the ICCPR with respect to free speech rights, and we cannot statistically distinguish the magnitudes of these two moderation effects.

This raises the question of why protecting the freedom of speech might be different from other rights in this context. One possible reason for this that we hope to analyze further in future work may be that free speech rights are defined more precisely in human rights treaties. The ICCPR and other human rights treaties do not define all protected rights with the same degree of precision.87 The freedom of speech stands out as a precisely defined right, which sets it apart from the freedoms of religion and of assembly and association. Unlike the others, the freedom of speech and expression spans two articles (Art. 19 and 20). More importantly, it is the only right for which the ICCPR does not define permissible limitations in generic terms,88 but instead provides specific and substantive constraints. For example, Art. 20 provides that war propaganda and “advocacy of national, racial or religious hatred that constitutes incitement to discrimination, hostility or violence” is not protected by the freedom of expression.

The precision of the freedom of speech may make it less costly to enforce using the direct effect mechanism than other rights. Precise rules limit the range of possible interpretations, thus reducing the likelihood of disputes over their interpretation.89 Judicial enforcement of more precise rights may therefore be less likely to provoke political backlash, because it leaves courts less vulnerable to criticism that they are ‘making’ law by interpreting human rights treaty provisions with indeterminate content. Thus, if and to the extent the precision of free speech rights reduces the probability of backlash, this would reduce the political risks of using the direct effects mechanism.

6. Conclusions

As Alter et al. observe in the introduction to this special issue, national courts play important roles in the judicialization of international relations. Among the most consistent arguments in the literature on international human rights treaties is that national courts can act as an enforcement mechanism, holding governments to their commitments and thus enhancing the effects of treaty membership, especially with respect to civil rights. Yet the existing literature has glossed over how domestic courts perform this enforcement function. As Alter et al. note, we need to better understand the processes by which courts conduct such adjudication.

87 We follow Abbott et al. (2000) in defining precision as the extent to which “rules unambiguously define the conduct they require, authorize, or proscribe” (p. 401).

88 By contrast, article 18.3, on the freedom of religion, states that “freedom to manifest one’s religion or beliefs may be subject only to such limitations as are prescribed by law and are necessary to protect public safety, order, health, or morals or the fundamental rights and freedoms of others.”

89 Lipson 1991. Versteeg and Zackin (forthcoming) likewise argue that more precise, or specific, constitutional rules reduce judicial discretion, and thus disputes over their interpretation.
This paper aims to contribute to our understanding of judicialization by examining the relationship between (a) the legal mechanisms courts can use to enforce human rights treaties and (b) the politics of human rights and judicial backlash. We focus on two legal mechanisms: direct effect and the canon of interpretation. These mechanisms differ not only in terms of the countries in which they are available to courts, but also in terms of the legal and political constraints courts face when using them. One of the key contributions of our paper is to explain these constraints and how they shape the relative effectiveness of enforcement mechanisms. We argue that, because of these constraints, the canon of interpretation is generally more effective, in the sense that its availability to domestic courts increases the effects of human rights treaty ratification more than the availability of the direct effect mechanism.

Our paper has implications for our understanding of international judicialization. First, our argument brings together the phases of the judicialization process. Our most direct focus is on the adjudication phase, in that we describe and assess the effectiveness of mechanisms available to courts when cases of human rights abuses arise. Yet, our focus is also on the compliance and feedback phases, insofar as our argument recognizes that strategic courts can anticipate reactions to their decisions, which influences their decision-making. Specifically, we argue that domestic courts’ use of legal mechanisms to enforce human rights treaties is constrained by the possibility of future backlash. In turn, this means our argument also has implications for the pre-litigation phase. Governments also act strategically and, to some extent, can be expected to understand the extent to which their decisions may be sanctioned by courts. Thus, the extent to which governments use the repressive tactics that lead to rights-claiming cases in the first place is influenced by the processes and mechanisms of both the litigation and post-litigation phases. Our argument therefore demonstrates one way in which the phases of the judicialization process are crucially dependent on each other.

Second, this paper enhances our understanding of international judicialization by examining in depth the role of national courts in this phenomenon. As Alter et al. note, “In countries in which ratified treaties are automatically part of the domestic legal order, courts have broad competence to review claims alleging violations of international law.” In addition, human rights is an issue-area in which this mechanism is often argued to be especially crucial because international enforcement is often costly or otherwise ineffective. Yet our paper demonstrates that, while national courts can enforce human rights treaties, their legal power can also be constrained by political concerns. Specifically, national courts may decide not to use their enforcement power when the legal mechanism available to them is more likely to generate political backlash. This implies that future work should analyze in depth both the legal and political concerns that can both facilitate and constrain the ability of domestic national courts to enforce international commitments.

This paper also has several implications regarding the effects of human rights treaties. First, as the literature has increasingly recognized, human rights treaty membership can reduce repression, but such effects are conditional on the availability of actors and institutions that can serve as enforcement mechanisms. Our argument furthers our understanding of these conditions by explaining how and under what conditions domestic courts can effectively enforce such treaties. Second, our argument has potential implications for our understanding of the
substitutability of repressive tactics. Governments have a repertoire of forms of repression at their disposal, and in choosing among these forms one factor they consider is the potential for legal sanctions. If and to the extent governments understand the relative effectiveness of judicial mechanisms, our argument implies that, in countries in which only the direct effect mechanism is available to courts, governments may be more likely to employ violations of the freedoms of association and religion than violations of the freedom of speech. More broadly, in countries in which the canon of interpretation is available to an independent domestic court, meaning that governments can expect civil rights violations to be sanctioned, the availability of this legal mechanism may lead to strategic substitution in favor of harsher tactics, such as physical integrity violations, which are more difficult for such courts to sanction. We hope to explore these possibility in future work.
Bibliography


Alter, Karen J., Emilie Hafner-Burton and Laurence R. Helfer. 2016. The Judicialization of International Relations. (Draft)


Hamilton, Alexander. 1788. The Federalist Papers, No. 78.


Staton, Jeffrey K. 2010. *Judicial Power and Strategic Communication in Mexico*. New York:


