

Human Rights Backsliding

Andrew T. Guzman* & Katerina Linos**

Human rights advocates and international lawyers view international agreements and other international norms as important tools to improve human rights around the world. This Article explains that, contrary to widely held beliefs, international human rights norms are not a one-way street. Norms capable of generating improved behavior in poorly performing states sometimes also exert a downward pull on high-performing states. This downward pull leads to what we term “human rights backsliding”—a tendency for high-performing states to weaken their domestic human rights regimes relative to prior behavior or relative to what they would otherwise have done.

The theory of backsliding is a novel one, so we introduce it with several real-world examples. In order to describe the theory, its assumptions, and its consequences as explicitly as possible, we also provide a formal model of backsliding. We then explain how an understanding of human rights backsliding helps explain state behavior that is otherwise puzzling. Finally, we explore some of the implications of backsliding for the design of international

Copyright © 2014 California Law Review, Inc. California Law Review, Inc. (CLR) is a California nonprofit corporation. CLR and the authors are solely responsible for the content of their publications.

* Jackson H. Ralston Professor of Law and Associate Dean for International and Advanced Degree Programs, University of California, Berkeley, School of Law. Harvard Law School, J.D.; Harvard University, Ph.D.

** Assistant Professor of Law, University of California, Berkeley, School of Law. Harvard Law School, J.D.; Harvard University, Ph.D. We are grateful to Daniel Abebe, Bill Alford, Roxanna Altholz, Robert Bartlett, Gabriella Blum, Anu Bradford, Bill Burke-White, David Caron, Robert Cooter, Jeffrey Dunoff, Stavros Gadinis, David Gartner, Tom Ginsburg, Jack Goldsmith, Ryan Goodman, Gillian Hadfield, Larry Helfer, Justin McCrary, Saira Mohamed, Mark Pollack, Eric Posner, Alison Post, Greg Shaffer, David Sklansky, David Sloss, Mila Versteeg, Mark Wu, and participants at the 2012 Berkeley Law Faculty Retreat, the Northern California International Law Workshop, the Junior International Law Scholars Association Meeting, the Cornell Law faculty workshop, the Temple International Law Workshop, the Harvard International Law Workshop, the American Society of International Law Research Forum, the Chicago International Law Workshop, the American Law and Economics Association Annual Meeting, and the Berkeley Law and Economics and Law and Society Joint Seminar for their helpful comments. Sylvia DeTar, Jerome Hsiang, Helena Jung, and May Whitaker provided excellent research assistance.

agreements, and we consider strategies for advocates seeking to advance human rights internationally.

Introduction	604
I. A Puzzle: How Does Backsliding Happen?	611
II. Three Examples of Backsliding.....	614
A. Defendants' Rights in the United Kingdom	615
B. Same-Sex Marriage in the United Kingdom and the United States.....	617
C. Women's Rights in Sweden.....	620
III. The Model	623
A. International Norm Higher than Domestic Equilibrium ($k > S^*$).....	628
B. International Norm Lower than Domestic Equilibrium ($k < S^*$).....	631
IV. Extending the Theory of Backsliding	635
A. Domestic Factors Influencing Backsliding.....	636
B. How the Content and Form of the Norm Matters.....	638
1. Floors, Ceilings, and Prohibitions.....	638
2. Hard Law Versus Soft Law.....	640
3. Multiple Standards	641
4. Vague Norms and Uncertainty.....	643
C. Other Theories of International Law's Influence	644
V. Implications.....	646
A. Treaty Membership.....	646
B. Some Empirical Puzzles Explained	648
1. The Behavior of High-Performing States	649
2. Regionalism.....	650
3. Aspirational Norms	652
Conclusion.....	654

INTRODUCTION

Human rights practices have improved dramatically in many parts of the world over the last century, in no small part because of the tireless efforts of transnational advocacy groups using international legal instruments.¹ A recent

1. See generally KATERINA LINOS, *THE DEMOCRATIC FOUNDATIONS OF POLICY DIFFUSION: HOW HEALTH, FAMILY AND EMPLOYMENT LAWS SPREAD ACROSS COUNTRIES* (2013); BETH A. SIMMONS, *MOBILIZING FOR HUMAN RIGHTS: INTERNATIONAL LAW IN DOMESTIC POLITICS* (2009); David Gartner, *Transnational Rights Enforcement*, 31 *BERKELEY J. INT'L L.* 1 (2013); Ryan Goodman & Derek Jinks, *How to Influence States: Socialization and International Human Rights Law*, 54 *DUKE L.J.* 621 (2004); Margaret E. Keck & Kathryn Sikkink, *Activists Beyond Borders: Advocacy Networks in International Politics*, in *EXPLORING INTERNATIONAL HUMAN RIGHTS* 98 (Rhonda L. Callaway & Julie Harrelson-Stephens eds., 1998) [hereinafter Keck & Sikkink, *Activists Beyond Borders*]; Margaret E. Keck & Kathryn Sikkink, *Transnational Advocacy Networks in International and Regional Politics*, 159 *INT'L SOC. SCI. J.* 89 (1999) [hereinafter Keck & Sikkink,

award-winning book by Professor Beth Simmons reports that after ratifying international human rights treaties, governments in many countries detain and torture fewer people,² allow for greater religious freedom,³ reduce child labor rates,⁴ and improve women's access to contraception.⁵ However, this same research notes that these effects are concentrated among moderately democratic countries and countries transitioning toward democracy.⁶ In contrast, treaty ratification often does not have these salutary effects in stable democracies.⁷ Most importantly for this Article, in stable democracies, ratification of international treaties is sometimes correlated with the erosion of human rights protections.⁸ These findings invite us to think carefully about the varied effects human rights treaties may have on state behavior.

In this Article, we question an implicit assumption held by many human rights scholars and advocates: that international human rights standards are a one-way street that can lead states only to expand domestic protections.⁹ Though international law and international norms can, indeed, be useful tools to improve human rights performance in poorly performing states, we argue that they can also undermine efforts to adopt or maintain high levels of protection in countries that would otherwise offer protections above the international norm. We call this phenomenon "human rights backsliding."

We define human rights backsliding as a process in which governments react to international standards by providing fewer or weaker human rights protections. Our definition includes the withdrawal of previously available rights as well as stasis or stagnation where we would otherwise expect to

Transnational Advocacy Networks]; Katerina Linos, *Diffusion Through Democracy*, 55 AM. J. POL. SCI. 678 (2011).

2. See SIMMONS, *supra* note 1, at 257.

3. *Id.* at 200.

4. *Id.* at 328.

5. *Id.* at 254.

6. *Id.* at 200, 254, 305, 328.

7. *Id.* at 200, 328; see also Emilie M. Hafner-Burton & Kiyoteru Tsutsui, *Human Rights in a Globalizing World: The Paradox of Empty Promises*, 110 AM. J. SOC. 1373, 1374–78 (2005) (suggesting that countries that ratify human rights agreements often do not improve their behavior); Oona A. Hathaway, *Do Human Rights Treaties Make a Difference?*, 111 YALE L.J. 1935, 1940–41 (2002) (same).

8. See SIMMONS, *supra* note 1, at 281–82 (suggesting that among high rule-of-law states, countries that have ratified the Convention Against Torture report a higher prevalence of torture than countries that have not ratified the Convention); Eric Neumayer, *Do International Human Rights Treaties Improve Respect for Human Rights?*, 49 J. CONFLICT RESOL. 925, 928–29 (2005) (exploring severe compliance challenges).

9. For related criticisms, see Eran Shor, *Conflict, Terrorism, and the Socialization of Human Rights Norms: The Spiral Model Revisited*, 55 SOC. PROBS. 117, 118 (2008) (critiquing prominent human rights scholars for their assumption that "once states adopt the rhetoric of human rights and begin to move towards norm compliance, there is no turning back. . . . [They] move forward uniformly toward norm compliance, or alternatively remain stagnant"); see also David Rieff, *The Precarious Triumph of Human Rights*, N.Y. TIMES MAG., Aug. 8, 1999, at 37 ("The human rights movement has assumed that establishing norms will lead to a better world.").

observe an expansion of rights (although it may be more difficult to clearly establish causation in the latter situation).

The notion of backsliding is novel, so there is some burden on us not only to explain that it is possible as a matter of theory, but also to show that it is plausible in fact. To this end, we consider the observations made by Professors Zachary Elkins, Thomas Ginsburg, and Beth Simmons regarding the Universal Declaration of Human Rights (UDHR).¹⁰ Following the adoption of the UDHR, some rights began to appear more frequently in national constitutions, including the right to life, the prohibition on ex post facto punishment, the right to join a trade union, the presumption of innocence in trial, the right to free movement, and the prohibition on cruel and inhuman treatment.¹¹ All of this is consistent with conventional views on how international human rights law can lead to the diffusion and expansion of rights domestically. Less consistent is the fact that certain other rights “may have fallen out of fashion in part due to their exclusion from the UDHR.”¹² The authors note, for example, that the right to a jury trial, the prohibition on censorship, the right to petition, certain intellectual property rights, prohibitions on child employment, and the right to a free press were less common in national constitutions after the UDHR was adopted than would otherwise be expected.¹³ These rights not only failed to advance following the UDHR, they became *less* popular than they would have been absent the UDHR’s adoption. This is an example of human rights backsliding at work.

In the pages that follow, we discuss additional examples of backsliding and focus on specific occurrences to illustrate the concept.¹⁴ First, we describe how the United Kingdom reduced the scope of criminal defendants’ rights to exclude hearsay evidence, influenced in part by decisions from the European Court of Human Rights (ECtHR) that failed to include the common law hearsay rule among the minimum protections required in all European countries.¹⁵ Second, we show how opponents of same-sex marriage in the United Kingdom used ECtHR decisions in their lobbying efforts to resist reform. The ECtHR held that the European Convention does not require states to legalize same-sex marriage, and conservative groups used these decisions to fight proposals to legalize same-sex marriage in the United Kingdom.¹⁶ Finally, we explain how Sweden responded to a European directive setting minimum standards relating to maternity-leave benefits by limiting the choices and benefits available to women. Swedish feminists, joined by advocates for

10. See Zachary Elkins, Tom Ginsburg & Beth A. Simmons, *Getting to Rights: Treaty Ratification, Constitutional Convergence, and Human Rights Practice*, 54 HARV. INT’L L.J. 61 (2013).

11. *Id.* at 81.

12. *Id.*

13. *Id.*

14. See discussion *infra* Part II.

15. See discussion *infra* Part II.A.

16. See discussion *infra* Part II.B.

children, the elderly, and the disabled, have expressed a concern that the European Union (EU) represents a threat to Sweden's generous welfare state.¹⁷ In other words, they feared human rights backsliding.

Developing a coherent theory of backsliding requires that we think seriously about how international norms affect domestic practices and that we make some assumptions about how those norms are transmitted to domestic policy makers. For most of this Article, we adopt a theory of domestic politics that is consistent with prevailing perspectives on how international human rights norms come to affect domestic policies.

According to major strands of the human rights literature, international standards can be effective because they focus attention on particular issues and place these issues on national agendas.¹⁸ In addition, international standards can influence the views of domestic publics, interest groups, and decision makers.¹⁹ When an international organization promotes a certain policy proposal as an international human right, this endorsement from a credible and disinterested outsider makes the proposal seem less radical and strengthens the rhetorical position of advocates for the position.²⁰ Rights that previously seemed undesirable or unattainable begin to appear feasible.²¹ At the margin, this “nudge” from an international institution can persuade citizens to support the policy, potentially tipping the balance in favor of the introduction of this right into the national legal system.²² At the core of our argument is the observation that, under reasonable assumptions, this same mechanism can also operate to lower human rights protections in high-performing states—or, in other words, to trigger human rights backsliding.

Statements from an international body can serve as a thumb on the scale and strengthen the position of both conservative and progressive groups. Protecting human rights inevitably involves some form of trade-off: granting criminal defendants more rights may limit victims' rights, wider access to food

17. See discussion *infra* Part II.C.

18. See SIMMONS, *supra* note 1, at 128–29.

19. See *id.* at 14–15; Goodman & Jinks, *supra* note 1, at 654–55; Keck & Sikkink, *Transnational Advocacy Networks*, *supra* note 1, at 89–90; Thomas Risse & Kathryn Sikkink, *The Socialization of International Human Rights Norms into Domestic Practices: Introduction*, in *THE POWER OF HUMAN RIGHTS: INTERNATIONAL NORMS AND DOMESTIC CHANGE* 1, 4–5, 16–17 (Thomas Risse, Stephen C. Ropp & Kathryn Sikkink eds., 1999). National standards may also serve to formalize emerging international human rights norms and clarify actual human rights behavior in a country. It may be that states are more likely to conform to prevailing norms when they have accurate information regarding other states' human rights behavior. For a discussion on the psychology of misperceiving a norm and its effects on incentives to comply with that norm, see Robert Cooter, Michal Feldman & Yuval Feldman, *The Misperception of Norms: The Psychology of Bias and the Economics of Equilibrium*, 4 *REV. L. & ECON.* 889, 891 (2008).

20. See LINOS, *supra* note 1, at 13–14, 28; Xinyuan Dai, *Why Comply? The Domestic Constituency Mechanism*, 59 *INT'L ORG.* 363, 363, 366, 384–85 (2005); Linos, *supra* note 1, at 679–80, 692.

21. See LINOS, *supra* note 1, 96–98.

22. *Id.* at 22–23.

and water may require higher taxes, and increased protection of free speech may require a relaxation of hate-speech codes. Domestic interest groups and political leaders line up on both sides of such debates. For example, in a country debating an expansion of domestic human rights protection above the international standard, opponents of the expansion can point to the international standard and argue that it reflects the right balance between diverse concerns. Protecting human rights any further, they might argue, would represent a radical and ill-thought-out experiment, compromise competing values (such as economic growth, law and order, national security, liberty, or any number of other priorities), and make their country worse off. Thus, just as international norms can lead to the expansion of rights in low-performing countries, a relatively low international standard can arm opponents of the expansion with a neutral, external benchmark and strengthen the persuasiveness of their arguments.²³

The above theory of human rights, in which domestic constituencies are essential to the transmission of international standards, is an important and conventional one, but it is not the only one. Toward the end of the Article, we consider other familiar theories of human rights transmission from international norm to domestic policy, and demonstrate that backsliding can result from any of these established mechanisms.²⁴

If we are correct that international human rights instruments not only expand human rights protections in poorly performing states, but also limit human rights practices among top performers, major implications follow for the design of international human rights regimes. There is some, albeit imperfect, evidence that drafters of major human rights instruments are worried about the possibility of backsliding. Many international agreements clearly state that they set minimum standards and that countries are free to set much higher goals. For example, the Convention on the Elimination of All Forms of Discrimination against Women (CEDAW) concludes by stating that “[n]othing in the present Convention shall affect any provisions that are more conducive to the achievement of equality between men and women which may be contained [in other national and international instruments].”²⁵ Additionally, both the International Covenant on Civil and Political Rights (ICCPR) and the International Covenant on Economic, Social and Cultural Rights (ICESCR) use substantively similar language to instruct that there shall be no restriction in national protections “on the pretext that the present Covenant does not

23. We do not claim that increases in human rights performance in low-performing countries and backsliding in high-performing countries are of the same magnitude. Our theory does, however, have a certain symmetry inasmuch as influences that increase (or decrease) the impact of international norms on low performers will also increase (or decrease) the impact on high performers, and vice versa.

24. See discussion *infra* Part IV.C.

25. Convention on the Elimination of All Forms of Discrimination Against Women art. 23, Dec. 18, 1979, 1249 U.N.T.S. 13.

recognize such rights or that it recognizes them to a lesser extent.”²⁶ Regional instruments also emphasize that the protections they offer should be understood as minima. For example, the European Convention on Human Rights (ECHR) concludes by specifying that “[n]othing in this Convention shall be construed as limiting or derogating from any of the human rights and fundamental freedoms which may be ensured under [other national and international instruments].”²⁷ Similarly, the American Convention of Human Rights highlights that it should not be read to “preclud[e] other rights or guarantees that are inherent in the human personality” or “restric[t] the enjoyment or exercise of any right or freedom recognized by virtue of the laws of any State Party.”²⁸ Indeed, this type of language is common in many treaties, and can be found both in general provisions like the ones just mentioned²⁹ and in the specification of the scope of particular rights.³⁰

This additional language is not offered simply to make the legal requirements clear, but also to provide emphasis. It signals the drafters’ concerns that domestic political actors might try to use the agreement to argue against higher standards. In other words, it is useful as a means to guard against backsliding.

26. International Covenant on Civil and Political Rights art. 5(2), Dec. 19, 1966, 999 U.N.T.S. 171; *see also* International Covenant on Economic, Social and Cultural Rights art. 5(2), Dec. 16, 1966, 993 U.N.T.S. 3 (“No restriction upon or derogation from any of the fundamental human rights recognized or existing in any country in virtue of law, conventions, regulations or custom shall be admitted on the pretext that the present Covenant does not recognize such rights or that it recognizes them to a lesser extent.”).

27. Convention for the Protection of Human Rights and Fundamental Freedoms art. 53, Nov. 4, 1950, 213 U.N.T.S. 221 [hereinafter ECHR].

28. American Convention on Human Rights: “Pact of San Jose, Costa Rica” art. 29, Nov. 22, 1969, 1144 U.N.T.S. 123 [hereinafter American Convention on Human Rights].

29. *See also* International Convention on the Protection of the Rights of All Migrant Workers and Members of Their Families art. 81, Dec. 18, 1990, 2220 U.N.T.S. 93 (“Nothing in the present Convention shall affect more favourable rights or freedoms granted to migrant workers and members of their families by virtue of: (a) The law or practice of a State Party; or (b) Any bilateral or multilateral treaty in force for the State Party concerned.”).

30. *See, e.g.*, International Convention for the Protection of All Persons from Enforced Disappearance, G.A. Res. 61/177, U.N. Doc. A/RES/61/177 (Dec. 20, 2006) (detailing the minimum scope of information that states must produce on persons deprived of liberty); Optional Protocol to the Convention on the Rights of the Child on the Sale of Children, Child Prostitution and Child Pornography art. 3, *opened for signature* May 25, 2000, 2171 U.N.T.S. 247 (calling on states to criminalize, at minimum, a list of offenses against children); International Convention on the Protection of the Rights of All Migrant Workers and Members of Their Families, *supra* note 29, art. 18 (articulating a set of minimum rights guarantees extended to migrant workers and family members when accused of crimes); American Convention on Human Rights, *supra* note 28, art. 8 (guaranteeing minimum rights guarantees for the criminally accused); International Covenant on Civil and Political Rights, *supra* note 26, art. 14(3) (outlining protections states must afford criminal defendants and explicitly providing that these are “minimum guarantees”); International Covenant on Economic, Social and Cultural Rights, *supra* note 26, art. 7(a) (describing how states must define, at a minimum, fair remuneration for workers); ECHR, *supra* note 27, art. 6 (specifically setting forth a list of minimum rights applicable to the accused).

While these drafting strategies are helpful, we do not believe that they are enough to fully prevent backsliding. Expressing a norm as a floor might be central to litigation surrounding human rights instruments, but it would not prevent backsliding as a result of the political mechanisms we consider. This is because backsliding is not required by the formal legal rules—it is an unintended consequence. In the discussion that follows, we consider this drafting question in detail along with related issues concerning the formal legal status of the norm, the possibility of a norm with multiple standards at different levels, and the consequences of vague norms.³¹

Understanding the risk of backsliding helps explain several patterns in international human rights law that have puzzled other theorists. One such puzzle is the reluctance of top-performing states to join international treaties whose standards they already meet or exceed. Another puzzle concerns the proliferation of regional human rights regimes. Regional regimes are surprising because human rights derive a good deal of their moral force from claims to universality: the notion that all human beings have fundamental rights by virtue of being human, not by virtue of residing in particular parts of the world. Finally, the risk of backsliding can help explain why human rights standards are often set at extremely high, and even unrealistic, levels, unlike other international treaty commitments that only call for modest reforms.³²

In sum, our contribution is threefold. First, we introduce the important possibility of backsliding to debates in international human rights law and challenge the implicit assumption that international agreements that specify floors serve only to improve human rights performance. We introduce our theoretical contribution in Part I and present some examples and illustrations of backsliding in Part II. Second, we develop a simple, formal model to clearly explain the mechanisms through which backsliding happens and the circumstances in which it is most likely to occur. Part III of our Article presents this formal model, and Part IV extends this model to discuss what changes when we vary some of the model's assumptions. The main contribution of the formal model is to clarify the logic of our argument and show precisely where we depart from existing theoretical writings. Third, we describe in some depth how the possibility of backsliding can explain major puzzles in international human rights law and outline important practical implications for states, NGOs, and others involved in the design of human rights instruments. These implications are presented in Part V. Finally, we conclude.

31. See discussion *infra* Part IV.B.

32. See discussion *infra* Part V.

I.

A PUZZLE: HOW DOES BACKSLIDING HAPPEN?

We begin by noting that recent evidence supports the view that human rights treaties can influence state behavior for the better, at least in some circumstances, and can lead to an expansion of domestic protections.³³ We highlight this point because our theory assumes that human rights instruments matter. Theories suggesting that human rights treaties have very little influence on state behavior rule out the possibility of backsliding because they exclude both positive and negative influences from abroad.³⁴

Empirical evidence establishing that human rights agreements can have a positive effect on state behavior cries out for a theoretical explanation. This evidence is surprising, at least initially, because some of the most prominent theories of international law do not work well in the human rights field. One such theory, institutionalism, treats states as rational and unitary actors that seek to maximize their own gains from the international system.³⁵ While this theory is helpful in explaining what we observe in many areas of international law, including the environment, trade, security, and more, it does not work well in the human rights field. Institutionalism performs poorly in this area because its assumption of selfish states cannot easily accommodate efforts to improve the well-being of human beings in other states. It cannot explain why an improvement in human rights abroad is valuable to the state.³⁶

Furthermore, even if one accepts that states wish to influence human rights policies abroad, why would the resulting treaties be effective? The familiar incentives to comply, known as the Three Rs of Compliance, are reciprocity, retaliation, and reputation.³⁷ These three incentives do not predict much compliance-pull in the human rights area.³⁸ When a state deprives its citizens of fundamental rights, other states are unlikely to reciprocate or

33. See SIMMONS, *supra* note 1, at 14–15.

34. See JACK L. GOLDSMITH & ERIC A. POSNER, *THE LIMITS OF INTERNATIONAL LAW* 108 (2005).

35. See ANDREW T. GUZMAN, *HOW INTERNATIONAL LAW WORKS: A RATIONAL CHOICE THEORY* 15–24 (2008) (discussing the relative advantages and disadvantages of analyzing international law through the lens of institutionalism).

36. See *id.* at 45 (explaining how reciprocity can fail as a way to induce compliance for human rights treaties) and 66–68 (explaining how retaliation may fail when trying to enforce multilateral human rights agreements).

37. *Id.*

38. See *id.* at 33–48; see also Andrew T. Guzman, *Formation of Human Rights Treaties, A Response to Alex Geisinger*, OPINIO JURIS (Feb. 11, 2008, 7:20 PM), <http://opiniojuris.org/2008/02/11/formation-of-human-rights-treaties-a-response-to-alex-geisinger/>; Andrew T. Guzman, *The Puzzle of Human Rights Treaties*, OPINIO JURIS (Feb. 12, 2008, 5:13 PM), <http://opiniojuris.org/2008/02/12/the-puzzle-of-human-rights-treaties/> (highlighting methodological difficulties with explaining human rights treaties).

retaliate by committing similar human rights violations.³⁹ Moreover, a state that engages in human rights violations may be able to develop separate reputations in other issue areas by continuing to comply with its trade, military, and other commitments.⁴⁰

If institutionalism fails, theorists and policymakers must seek a different approach. In particular, at least one of the assumptions of institutionalism must be relaxed. We choose to relax the unitary state assumption. This is the assumption that states are unified entities that seek to maximize a well-defined national interest. Relaxing the unitary state assumption has the obvious appeal of eliminating one of the least plausible assumptions of the classic rational choice model. Opening the black box of the state, however, poses myriad challenges. It is one thing to observe that domestic politics matters, but quite another to describe a model of domestic politics that is both realistic and tractable.

In the paragraphs that follow, we outline such a model with a focus on how government leaders respond to pressures from voters and interest groups. The claim that government leaders respond to domestic political pressure is, of course, not novel. Indeed, we believe it is the most plausible mechanism available to explain why international human rights instruments often improve state behavior. We take this well-established theory and develop an important implication that has not been highlighted to date—the possibility of backsliding.

International agreements can influence domestic human rights debates because human rights policies involve important trade-offs and place different societal groups in opposing positions. For example, an investigation into allegations of police misconduct may please human rights advocates seeking justice and accountability, but displease police forces and hobble security efforts. A law increasing minimum labor standards may pit labor unions seeking improved working conditions against employer associations concerned about competitiveness and economic growth. Expanded access to water and sanitation may help poor communities enjoy fundamental social and economic rights, but may involve cuts in other government programs and trigger significant criticism from these programs' beneficiaries.

In making these trade-offs, politicians seek to maintain the support of the public at large. Popular support is critical to democratic leaders concerned about the next election,⁴¹ and even for autocrats, social unrest is highly

39. There are some exceptions. If a particular ethnic group controls a state, for example, it may respond to foreign conduct that harms members of the same ethnic group abroad with retaliatory actions against a minority ethnic group within its borders.

40. See GUZMAN, *supra* note 35, at 100–06; Rachel Brewster, *The Limits of Reputation on Compliance*, 1 INT'L THEORY 323, 327–28 (2009).

41. See LINOS, *supra* note 1, at 1–12, 19, 30; see also Donald Wittman, *Candidate Motivation: A Synthesis of Alternative Theories*, 77 AM. POL. SCI. REV. 142, 142 (1983).

problematic.⁴² Politicians who take radical positions that please special interest groups alone can lose voter support. Thus, politicians try to adopt relatively mainstream policies that are consistent with voters' values and likely to benefit the country as a whole. Consequently, both politicians and interest groups often try to present their positions in ways that appear mainstream and beneficial to the public at large.⁴³

International law, international norms, and the practices of other countries can serve as benchmarks against which citizens evaluate government performance. In deciding between a new, untested policy, and a policy that is endorsed by international organizations and widely adopted around the world, governments may often select the latter because it is easier to justify to domestic audiences.⁴⁴

For example, imagine that labor unions in a developing country seek to eliminate child labor in order to increase wages for adult workers and reduce accidents. If international law prohibits child labor, and many neighboring developing countries have restricted the practice, labor unions will have an easier time mobilizing widespread support for their position. As Beth Simmons puts it, "local agents have the motive to use whatever tools may be available and potentially effective to further rights from which they think they may benefit."⁴⁵ International law, international norms, and foreign states' practices can be powerful tools that are often referenced in pursuit of diverse domestic objectives.⁴⁶

All else being equal, conformity with an international norm gives a government greater political support at home and, in this sense, is valuable to leaders. The heart of the theory, then, is that an international signal regarding "proper" or "expected" human rights conduct will empower local interest groups, giving them a domestic political advantage and drawing local human rights policy closer to the standard specified in the international agreement or other instrument. The international focal point generates a gravitational pull on policy. This theory explains why an international agreement without strong enforcement provisions can nevertheless impact human rights outcomes.⁴⁷

42. See Jessica L. Weeks, *Autocratic Audience Costs: Regime Type and Signaling Resolve*, 62 INT'L ORG. 35, 62 (2008).

43. See, e.g., Emilie M. Hafner-Burton, Laurence R. Helfer & Christopher J. Farris, *Emergency and Escape: Explaining Derogations from Human Rights Treaties*, 65 INT'L ORG. 673 (2011) (arguing that governments use derogation clauses in human rights treaties to suspend their obligations in time of emergency and to assure domestic audiences that the emergency rights restrictions are necessary, legal, and temporary deviations from normal levels of rights protections).

44. See LINOS, *supra* note 1, at 2–6; Linos, *supra* note 1.

45. SIMMONS, *supra* note 1, at 373.

46. See generally Keck & Sikkink, *Activists Beyond Borders*, *supra* note 1; Keck & Sikkink, *Transnational Advocacy Networks*, *supra* note 1; Risse & Sikkink, *supra* note 19.

47. This argument draws most directly on the work of Katerina Linos, but is consistent with arguments made by many other authors, to explain why human rights instruments influence state behavior. See generally LINOS, *supra* note 1, at 175–85; SIMMONS, *supra* note 1; Goodman & Jinks,

We demonstrate that this same mechanism can also limit the scope of particular rights. Imagine that instead of banning child labor, international agreements take a more nuanced position and argue that “children’s or adolescents’ participation in work that does not affect their health and personal development or interfere with their schooling, is generally regarded as being something positive.”⁴⁸ This is in fact the International Labour Organization’s (ILO) current position.⁴⁹ In this situation, employer associations and conservative politicians can point to the ILO’s definition in their efforts to allow certain types of child labor and to fight labor unions’ efforts to completely ban the practice.

The argument we propose—that international norms can lead some countries to offer lower protections than they otherwise would—works best under certain conditions. First, backsliding is most likely to occur in cases in which domestic leaders are responsive to the public at large. This is not a major scope limitation for our argument because the states that offer high levels of human rights protections, and thus that could potentially be dragged down by international human rights standards, are very likely to be representative democracies. Second, we assume, as most of the human rights literature does, that states are responsive to human rights norms. Again, this assumption is particularly plausible for the subset of countries at risk of backsliding. Democracies that offer unusually high levels of human rights protections tend to be among the most fervent supporters of international law and international norms, and are often highly integrated in transnational networks. Third, we assume that in many areas international standards are set at moderate levels. If international agreements instead set maximal standards that no government is likely to fully meet, this would eliminate the risk of backsliding. As we explain below, maximal standards would also limit the possibility of positively influencing poor performers.⁵⁰ Before presenting our formal model in Part III, we offer three examples of backsliding at work.

II.

THREE EXAMPLES OF BACKSLIDING

This Section provides some examples of backsliding in action. We explore how defendants’ rights declined in the United Kingdom in part because of minimum European standards, how same-sex marriage opponents used European Court decisions to fight the legalization of same-sex marriage, and

supra note 1; Linos, *supra* note 1; Keck & Sikkink, *Activists Beyond Borders*, *supra* note 1; Keck & Sikkink, *Transnational Advocacy Networks*, *supra* note 1; Risse & Sikkink, *supra* note 19; Dai, *supra* note 20.

48. *What is Child Labour*, INTERNATIONAL LABOUR ORGANIZATION, <http://www.ilo.org/ipec/facts/lang--en/index.htm> (last visited Feb. 25, 2014).

49. *Id.*

50. Such agreements exist, and we suggest in Part V.B.3 that concerns about backsliding may explain why they are written with such unrealistically high requirements.

why women's rights advocates in Sweden saw European maternity and parental-leave minimums as a threat.

Each of our examples involves clear tradeoffs, as we believe that the protection of human rights typically requires governments to balance competing concerns. One could instead argue that the prohibitions on torture, slavery, or genocide should function as absolute prohibitions. Yet, as we explain in Part IV below, even prohibitions on such grave violations involve trade-offs, as it is possible to define each of these practices broadly or narrowly. In addition, we highlight that our cases come disproportionately from Europe. As we explain in Part IV below, this is not a coincidence—the risk of backsliding is highest in democratic states that offer high levels of rights protections and are very integrated in the international community.⁵¹ In contrast, countries such as the United States that are less integrated into the international community and resist increasing rights protections as a result of international standards may face lower risks of backsliding.⁵²

A. Defendants' Rights in the United Kingdom

In the United States, the common law rule against hearsay evidence—in conjunction with the constitutional right to confrontation—offers criminal defendants important protections by disallowing the admission of out-of-court statements from persons who are not testifying at trial.⁵³ But the United Kingdom, the country that first developed the common law hearsay rule, has drastically reduced the rights of criminal defendants to question the prosecution's witnesses.⁵⁴ Specifically, reforms introduced in the 2003 Criminal Justice Act allow the prosecution to introduce hearsay evidence in the form of pretrial statements from persons unavailable to testify at trial.⁵⁵

The European Convention on Human Rights sets minimum human rights standards for dozens of European states, including the United Kingdom.⁵⁶ Article 6(3)(d) of the Convention states that every person charged with a crime

51. See discussion *infra* Part IV.A.

52. See *id.*

53. See Note, William Worthington, *Hearsay and Confrontation: Can the Criminal Defendant's Rights Be Preserved Under a Bifurcated Standard?*, 32 WASH. & LEE L. REV. 243, 243 (1975).

54. See David Alan Sklansky, *Hearsay's Last Hurrah*, 2009 SUP. CT. REV. 1, 2 (noting that the United Kingdom, along with several other common law countries, has reduced the scope of hearsay protections).

55. See Criminal Justice Act, 2003, c. 44, § 116 (1)–(2) (U.K.) (allowing the testimony of witnesses unavailable at trial due to death, physical or mental unfitness, absence from the United Kingdom when a return is not practical, disappearance despite reasonable efforts to locate the witness, or fear that the witness testimony will lead to recrimination); see also Sklansky, *supra* note 54, at 28–29.

56. See Luzius Wildhaber, *The European Court of Human Rights: The Past, the Present, the Future*, 22 AM. U. INT'L L. REV. 521, 526, 537–38 (2007) (essay by the President of the European Court of Human Rights from 1997 to 2007, explaining the role of the Court).

has the right “to examine or have examined witnesses against him.”⁵⁷ However, “as construed by the ECHR, the opportunity to cross-examine adverse witnesses is not invariably mandated by Article 6.”⁵⁸ For example, in *Isgrò v. Italy*, the ECtHR unanimously upheld Mr. Isgrò’s conviction even though Mr. Isgrò’s attorney was not able to cross-examine Mr. Isgrò’s primary accuser.⁵⁹

The Criminal Justice Act was largely based on a 1997 Law Commission Report titled *Evidence in Criminal Proceedings: Hearsay and Related Topics*.⁶⁰ This report offers several reasons for why the United Kingdom should limit the scope of its hearsay rule, and discusses the European Convention, the jurisprudence of the European Court, and the practices of continental European countries extensively.⁶¹ Additionally, the Commission’s report finds it “significant that in many European countries what would be called hearsay in England and Wales is admissible, and does not appear to be in contravention of the [European] Convention.”⁶² The Commission’s report concludes, “[T]he Convention does not require direct supporting evidence where it is sought to prove a particular element of the offence by hearsay. Adequate protection for the accused will be provided by the safeguards we propose.”⁶³ In short, while the United Kingdom likely had multiple reasons to reduce the rights of criminal defendants, the fact that it was out of line with the rest of Europe appears to have been part of the decision process.⁶⁴

Neither the European Convention nor the resulting jurisprudence was ever intended to reduce criminal defendants’ rights (or any other rights). Rather, its goal was to secure a minimum set of shared protections across Europe.⁶⁵ Nevertheless, influenced at least in part by these minimum standards, the

57. See ECHR, *supra* note 27, at art. 6(3)(d).

58. Todd E. Pettys, *Counsel and Confrontation*, 94 MINN. L. REV. 201, 247 (2009).

59. *Isgrò v. Italy*, 194 Eur. Ct. H.R. (ser. A) at 2, 7–10 (1991).

60. See ADRIAN KEANE, *THE MODERN LAW OF EVIDENCE 270* (7th ed. 2008); see also *R v. Horncastle*, [2009] UKSC 14, [2010] 2 A.C. 373, ¶ 29 (appeal taken from Eng.).

61. THE LAW COMMISSION, *EVIDENCE IN CRIMINAL PROCEEDINGS: HEARSAY AND RELATED TOPICS 56–57* (1997), available at http://lawcommission.justice.gov.uk/docs/lc245_evidence_in_criminal_proceedings_hearsay_and_related_topics.pdf.

62. *Id.* at 66; see also *id.* at 57 (explaining that practices that “fall foul of the hearsay rule in England and Wales . . . would be considered unobjectionable in most Continental systems” and might not violate the European convention, “depend[ing] on all the circumstances taken together”).

63. *Id.* at 67.

64. There has been a largely concurrent trend throughout the common law world to weaken the hearsay rule. Moreover, most of these jurisdictions, including Canada, Australia, and New Zealand, are outside the formal purview of the ECtHR, suggesting that common law hearsay reform may have occurred for a variety of reasons. Nonetheless, we believe that, at least in the case of the United Kingdom, the well-documented conversation between the ECtHR and both the British courts and legal commentators provides evidence that the ECtHR provided a significant “nudge” that allowed the U.K. to proceed with hearsay reform.

65. See Wildhaber, *supra* note 56.

United Kingdom reduced its criminal protections. It experienced human rights backsliding.

B. Same-Sex Marriage in the United Kingdom and the United States

Our definition of backsliding includes not only a withdrawal of previously available rights, but also use of international norms to limit or delay domestic efforts to expand rights. While England and Wales recently introduced same-sex marriage legislation, it is important to note that ECtHR decisions were used to bolster opposition to this reform. The relatively low international standard armed opponents of the expansion with an apparently neutral external benchmark and strengthened the persuasiveness of their arguments vis-à-vis undecided citizens.

Since the ECtHR's 1981 decision decriminalizing sodomy in *Dudgeon v. United Kingdom*,⁶⁶ some regional and international bodies have offered greater protection to gays and lesbians than many national legal systems. Advocates seeking to improve the status of gays and lesbians within domestic systems have responded by adopting the language of international human rights to persuade national decision makers and the national public that their country is out of step with the world.⁶⁷ This framing based on international human rights has, to some extent, replaced alternative framings, such as "national civil rights conceptions" and "framings based on gay and lesbian liberation and emancipation."⁶⁸

ECtHR decisions, however, do not always go as far as progressives might like. Relatively conservative decisions can undermine efforts to expand human rights in countries that would otherwise offer protections beyond the international norm. In two recent decisions, *Schalk and Kopf v. Austria*⁶⁹ and *Gas and Dubois v. France*,⁷⁰ the ECtHR held that the European Convention on Human Rights does not require member states' governments to grant same-sex couples access to marriage. These decisions were influential in the United Kingdom, a high performer in many areas of human rights. In the United

66. *Dudgeon v. United Kingdom*, 45 Eur. Ct. H.R. (ser. A) (1981).

67. See Laurence R. Helfer & Erik Voeten, *International Courts as Agents of Legal Change: Evidence from LGBT Rights in Europe*, 68 INT'L ORG. 77 (2014). Gay rights advocates point to ECtHR decisions such as *Dudgeon v. UK*, to UN Human Rights Committee Decisions such as *Toonen v. Australia*, to UN Human Rights Council resolutions condemning discrimination and violence against LGBT people, and even to statements by UN Secretaries-General. Human Rights Committee, *Toonen v. Australia*, Comm. No. 488/1992, U.N. Doc CCPR/C/50/D/488/1992 (Mar. 31, 1994), available at <http://www1.umn.edu/humanrts/undocs/html/vws488.htm>; Human Rights Council Res. 17/19, Rep. of the Human Rights Council, 17th Sess., May 30–June 17, 2011, A/HRC/17/L.9/Rev. 1 (June 17, 2011); Ronald Holzacker, "Gay Rights are Human Rights": *The Framing of New Interpretations of International Human Rights Norms* 1, 13–20 (Am. Political Sci. Ass'n Annual Meeting Paper, 2011), available at http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1902165.

68. Holzacker, *supra* note 67, at 1.

69. *Schalk and Kopf v. Austria*, 2010 Eur. Ct. H.R.

70. *Gas and Dubois v. France*, 2012 Eur. Ct. H.R.

Kingdom, domestic support for same-sex marriage has been widespread, as both the opposition Labour Party and the Conservative government of David Cameron, in coalition with the Liberal Democrats, have argued for this reform.⁷¹ Opponents of same-sex marriage, however, have strengthened their opposition by taking a page from the international human rights strategy of gay rights advocates and using the ECtHR rulings to support their arguments.⁷² The Church of England, for example, emphasized the ECtHR's rulings in explaining its opposition to same-sex marriage.⁷³ Similarly, conservative advocacy groups have used the ECtHR rulings to argue against same-sex marriage. For instance, Norman Wells from the Family Education Trust said “[t]he ruling from the ECHR will embolden those whose concerns about same-sex marriage and adoption are not inspired by personal hatred and animosity, but by a genuine concern for the well-being of children and the welfare of society.”⁷⁴

The ECtHR decisions also found their way into the debate through the media.⁷⁵ The rulings themselves are nuanced and leave room for a diversity of national positions on the issue, but they are often conveyed to the public in much simpler, conclusory terms. Many press accounts simply reported that the ECtHR had rejected the notion of same-sex marriage as a human right. *The Telegraph* headline was “Gay marriage is not a human right, according to

71. See generally David Cameron, Prime Minister of the United Kingdom, Keynote Address at the Conservative Party Conference (Oct. 5, 2011) (transcript available at <http://www.guardian.co.uk/politics/2011/oct/05/david-cameron-conservative-party-speech>) (“So I don’t support gay marriage despite being a Conservative. I support gay marriage because I’m a Conservative.”); *The Government Should Go Further Than They Currently Plan on Same Sex Marriage—Cooper*, LABOUR.ORG (Mar. 15, 2012), <http://www.labour.org.uk/government-should-go-further-on-same-sex-marriage> (a statement by Yvette Cooper, Labour’s Shadow Home Secretary and Shadow Minister for Women & Equalities, calling on the U.K. government to go further than what was currently proposed on gay marriage). The United Kingdom Parliament has since passed the Marriage (Same Sex Couples) Act 2013 to legalize gay marriage.

72. Amicus Curiae Brief of International Jurists & Academics as Amicus Curiae Supporting Petitioner Hollingsworth and Respondent Bipartisan Legal Advisory Group Addressing the Merits & Supporting Reversal at 24–26, *Hollingsworth v. Perry*, 133 S. Ct. 2652 (2013) (No. 12-144), *United States v. Windsor*, 133 S. Ct. 2675 (2013) (No. 12-307).

73. See A RESPONSE TO THE GOVERNMENT EQUALITIES OFFICE CONSULTATION—“EQUAL CIVIL MARRIAGE”—FROM THE CHURCH OF ENGLAND 10–12 (2012), available at <http://www.churchofengland.org/media/1475149/s-s%20marriage.pdf>.

74. Steve Doughty, *Gay Marriage Is not a “Human Right”: European Ruling Torpedoes Coalition Stance*, THE DAILY MAIL (Mar. 20, 2012), <http://www.dailymail.co.uk/news/article-2117920/Gay-marriage-human-right-European-ruling-torpedoes-Coalition-stance.html>; see also SBrinkmann, *European Court Rules Same-Sex Marriage Is Not a Human Right*, WOMEN OF GRACE BLOG (Mar. 27, 2012), <http://www.womenofgrace.com/blog/?p=13371>. However, same-sex marriage proponents countered references to *Schalk and Kopf v. Austria* and *Gas and Dubois v. France* by using examples from other European countries that had previously recognized same-sex unions. See also LINOS, *supra* note 1, at 33 (explaining how international and comparative benchmarks may pull in different directions, and why a single model may be helpful in defining a European norm).

75. See, e.g., Donna Bowater, *Gay Marriage Is Not a Human Right, According to European Ruling*, THE TELEGRAPH (Mar. 21, 2012), <http://www.telegraph.co.uk/news/religion/9157029/Gay-marriage-is-not-a-human-right-according-to-European-ruling.html>.

European ruling.”⁷⁶ The *Daily Mail* declared, “Gay marriage is not a ‘human right’: European ruling torpedoed Coalition stance.”⁷⁷

Indeed, both proponents and opponents of same-sex marriage in the United States are taking a page from the same playbook. For example, international and comparative law experts filed an amicus brief in the Supreme Court case *Hollingsworth v. Perry*⁷⁸ in support of respondents seeking marriage equality. They provided examples of how international and foreign courts have upheld fundamental notions of equal protection, liberty, and dignity, and have not contravened religious freedom.⁷⁹ But in a less expected move, prominent conservative opponents of same-sex marriage are also now invoking international law to buttress their argument that same-sex marriage is not an international norm. A number of conservative-leaning legal commentators also filed an amicus brief in *Hollingsworth*, as well as *United States v. Windsor*,⁸⁰ arguing that the ECtHR “found no such consensus [regarding same-sex marriage] in Europe.”⁸¹ They also cited the ECtHR in *Schalk and Kopf v. Austria*⁸² to show that Article 12 of the European Convention on Human Rights “does not impose an obligation on the respondent Government to grant a same-sex couple like the applicants access to marriage.”⁸³ In his dissent in *Windsor*, Justice Alito, joined by Justice Thomas, argued that “the right to same-sex marriage [is not] deeply rooted in the traditions of other nations. No country allowed same-sex couples to marry until the Netherlands did so in 2000.”⁸⁴ Despite historical discomfort among conservatives with the practice of domestic invocations of international law, Alito was the only Justice to cite to foreign practice in both *Hollingsworth* and *Windsor*, and he did so with the purpose of countering the Supreme Court’s tentative support of same-sex equality.

76. *Id.*

77. Doughty, *supra* note 74.

78. 133 S. Ct. 2652 (2013).

79. Brief for Foreign & Comparative Law Experts Harold Hongju Koh et al. as Amici Curiae Supporting Respondents at 5, 12, 38, *Hollingsworth v. Perry*, 133 S. Ct. 2652 (2013) (No. 12-144).

80. *Perry*, 133 S. Ct. at 2675.

81. Brief for International Jurists & Academics, *supra* note 72, at 25.

82. *Schalk and Kopf v. Austria*, 2010 Eur. Ct. H.R.

83. *Id.* at 26.

84. *United States v. Windsor*, 133 S. Ct. at 2715 (Alito, J., dissenting) (citing Ian Curry-Sumner, *A Patchwork of Partnerships: Comparative Overview of Registration Schemes in Europe*, in LEGAL RECOGNITION OF SAME-SEX PARTNERSHIPS IN EUROPE 71, 72 (Katharina Boele-Woelki & Angelika Fuchs eds., rev. 2d ed., 2012)). Justice Alito also notes that “virtually every culture, including many not influenced by the Abrahamic religions, has limited marriage to people of the opposite sex.” *Id.* at 2718 (citing Brief on the Merits for Respondent Bipartisan Legal Advisory Group of the U.S. House of Representatives at 2, *United States v. Windsor*, 133 S. Ct. 2675 (2013) (No. 12-307) (“Until a few decades ago, it was an accepted truth for almost everyone who ever lived, in any society in which marriage existed, that there could be marriages only between participants of different sex.”)).

We highlight the examples of same-sex marriage debates in the United Kingdom and the United States to show how relatively low international standards can empower opponents of same-sex marriage, just as relatively high international standards in the past have empowered gay rights advocates. We claim that the exclusion of same-sex marriage from a transnational set of human rights standards made domestic advocacy for same-sex marriage harder than it would otherwise have been. All else being equal, gay rights advocates would prefer that national publics not be exposed to headlines that “gay marriage is not a human right.” That said, low international standards on their own do not suffice to block progressive reforms, just as high international standards do not suffice to induce them. International standards are one among several factors that influence domestic policy decisions.

C. Women’s Rights in Sweden

Our third example of backsliding illustrates a reduction in parental-leave benefits in Sweden following a European directive intended to establish minimum standards in this area. In this particular case, concerns about backsliding were strong enough to provoke opposition to the entire project of European integration from some advocates of the Swedish welfare state.

Many people value parental leave as an important vehicle to allow working parents, and especially working women, to continue with their careers. By some estimates, maternity-leave laws can increase the employment rate of women of childbearing age (25–34) by 7 to 9 percent.⁸⁵ On the other hand, parental leave imposes regulatory burdens on employers and restricts the freedom of businesses to operate as they wish. Furthermore, a more generous leave policy confers benefits only on parents who chose to work rather than stay at home.⁸⁶ In short, parental-leave policies, like all policies, are subject to political contestation.

In the course of domestic debates on parental leave, international standards can be used by both progressives and conservatives. International norms that are higher than existing or proposed domestic norms can strengthen the position of those seeking more protective parental-leave policies, whereas international norms that are lower than domestic norms can strengthen the position of those supporting less protective parental-leave policies.

85. See Christopher J. Ruhm, *The Economic Consequences of Parental Leave Mandates: Lessons from Europe*, 113 Q. J. ECON. 285, 310–11 (1998); see also CATHERINE R. ALBISTON, INSTITUTIONAL INEQUALITY AND THE MOBILIZATION OF THE FAMILY AND MEDICAL LEAVE ACT: RIGHTS ON LEAVE (2010) (explaining how lack of information and uneven enforcement of leave laws influence take up); Christine Jolls, *Accommodation Mandates*, 53 STAN. L. REV. 223, 296–300 (2000) (discussing the effects of the FMLA); Gillian Lester, *A Defense of Paid Family Leave*, 28 HARV. J.L. & GENDER 1 (2005) (reviewing empirical work concluding that paid leave increases women’s labor force participation, and arguing that this is normatively desirable).

86. See generally Lester, *supra* note 85 (reviewing the arguments in favor and against maternity leave); Linos, *supra* note 1 (explaining the political economy of maternity leave debates).

Sweden is a global leader with respect to expanded parental leave. Starting in 1974, Sweden began offering six months of leave to either parent, compensated at 90 percent of their prior salary.⁸⁷ However, following the introduction of minimum standards across the European Union, Sweden cut back its generous benefits and limited women's choices.⁸⁸

Shortly after Sweden started seriously debating entrance to the EU in 1989,⁸⁹ and at least until Sweden ultimately joined the EU in 1995, several defenders of the generous Swedish welfare state expressed concerns about whether EU membership would lessen protections in Sweden.⁹⁰ Among their worries was a European directive setting minimum maternity leave standards throughout the EU.⁹¹ The EU Pregnant Workers directive, formally proposed by the EU Commission in 1990, and ultimately adopted in 1992, specified that all EU member states must offer a minimum of fourteen weeks of maternity leave.⁹² Moreover, the directive specified that this leave should be adequately compensated,⁹³ and that compensation comparable to that provided for sick leave is considered adequate.⁹⁴ In addition, the directive specified that all women should be obligated to take at least two weeks of leave.⁹⁵ This mandatory two-week minimum reflected concerns that some women might face employer pressure to turn down the leave. Both the preamble of the directive and its text clearly indicate that the directive is intended to set

87. See Ann-Zofie Duvander, Tommy Ferrarini & Sara Thalberg, *Swedish Parental Leave and Gender Equality: Achievements and Reform Challenges in a European Perspective*, 11 INST. FOR FUTURES STUD. 3, 8 (2005).

88. This discussion focuses on reforms in the early and mid-1990s, a period during which the EU promulgated a major directive on minimum maternity-leave standards, and a period during which Sweden joined the European Union. There have since been additional reforms to Swedish parental-leave benefits. For an overview of current benefits, see 13 ch. 10–14 §§ Socialförsäkringsbalk (Svensk författningssamling [SFS] 2010:110) (Swed.), available at http://www.riksdagen.se/sv/Dokument-Lagar/Lagar/Svenskforfattningssamling/Socialforsakringsbalk-201011_sfs-2010-110/?bet=2010:110#K13. We are very grateful to Helena Jung for locating and translating this and other Swedish sources for us.

89. *Sweden's Road to EU membership*, GOVERNMENT OFFICES OF SWEDEN, <http://www.government.se/sb/d/3470/a/20685>. Sweden eventually joined the EU in 1995 (last visited Feb. 25, 2014).

90. For many examples and a summary of these concerns, see MARIKA EHRENKRONA, GRANSKNING AV EU-KRITIKEN 45–51 (2001). To translate one of the many examples listed by Ehrenkrona: “The handicapped, children, the old and the sick will be abandoned. The nurturing model has come to an end. . . . What our fathers toiled to create, the EU will take from us.” *Id.* at 50–51.

91. See Milena Sunnus, *EU Challenges to the Pioneer in Gender Equality: The Case of Sweden*, in GENDERING EUROPEANISATION 223, 239, 240–45 (Ulrike Liebert ed., 2003); see also EHRENKRONA, *supra* note 90 (giving an overview of what the EU-critical voices in Sweden had to say when Sweden was about to have a referendum on EU membership).

92. Council Directive (EEC) No. 92/85 of 19 Oct. 1992, art. 8, 1992 O.J. (L 348) 1, available at <http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=CELEX:31992L0085:EN:NOT>. For a legislative history of the directive, see *Proposal for a Council Directive Concerning the Protection at Work of Pregnant Women or Women Who Have Recently Given Birth*, PRELEX, http://ec.europa.eu/prelex/detail_dossier_real.cfm?CL=en&DosId=105378#124443 (last visited Feb. 26, 2014).

93. Council Directive No. 92/85, *supra* note 92, at art. 11, § 2.

94. *Id.* at art. 11, § 3.

95. *Id.* at art. 8.

minimum requirements and that EU member states should not reduce existing protections if they are higher than the directive's minimum requirements.⁹⁶

At first, Swedish objections focused on the two weeks of mandatory leave. Swedish feminists saw this element as a stereotypical allocation of parental responsibilities to the female parent, and as a restriction on women's choices. The Swedish government took these concerns very seriously. Whereas Sweden typically transposes EU directives very promptly, Sweden delayed the introduction of this provision of the directive, and only transposed it years later, after the EU Commission had started infringement proceedings.⁹⁷

A larger concern arose from the fact that Swedish legislation transposing the EU directive also reduced the compensation levels for maternity leave to the levels of compensation typical of sick leave. Previously, parental leave had been compensated at 90 percent of a worker's salary, but the Swedish legislation transposing the EU directive reduced compensation to 80 percent, and later legislation reduced this even further to 75 percent.⁹⁸ The Swedish government's report justifying this reduction explicitly referenced the European standards and stated that maternity leave payments that were in line with sick leave payments were adequate.⁹⁹

To be clear, the policy choices made by Sweden were the result of many factors, as is always the case with significant political decisions. That said, the evidence suggests that European minimum standards put some downward pressure on domestic policies. Opponents of the policy change certainly seemed to believe as much: "[A]fter Sweden's entry in the EU, women experienced reductions of parental-leave payments from 90 p.c. to 75 p.c. of their income."¹⁰⁰

96. See, e.g., *id.* at pmbl. ("Whereas Article 118a of the Treaty provides that the Council shall adopt, by means of directives, minimum requirements for encouraging improvements, especially in the working environment, to protect the safety and health of workers; Whereas this Directive does not justify any reduction in levels of protection already achieved in individual Member States."); see also *id.* at art. 1, § 3 ("This Directive may not have the effect of reducing the level of protection afforded to pregnant workers . . . as compared with the situation which exists in each Member State on the date on which this Directive is adopted.").

97. See Gerda Falkner et al., *Non-Compliance with EU Directives in the Member States: Opposition Through the Backdoor?*, 27 W. EUR. POL. 452, 458–59 (2004).

98. See 28 ch. 7, 12 §§, 12 ch. 18 § Socialförsäkringsbalk (Svensk författningssamling [SFS] 2010:110) (Swed.), available at http://www.riksdagen.se/sv/Dokument-Lagar/Lagar/Svenskforfattningssamling/Socialforsakringsbalk-201011_sfs-2010-110/?bet=2010:110. (translated from Swedish by Helena Jung).

99. Proposition [Prop.] 1994/1995:207 Ny föräldraledighetslag m.m. [government bill] (Swed.), at ch. 4.8, available at http://www.riksdagen.se/sv/Dokument-Lagar/Forslag/Propositioner-och-skrivelser/Ny-foraldraledighetslag-mm_GI03207/?html=true (translated from Swedish by Helena Jung).

100. Ulrike Liebert, *Constructing Monetary Union: Euro-Scepticism, and the Emerging European Public Space* 16 (2001) (unpublished manuscript) (on file with authors) (describing how feminists have been critical of EU measures).

Concern about backsliding was not limited to parental leave. Activists also expressed concerns about Europe's potential influence on Sweden's generous benefits to children, the elderly, and the disabled.¹⁰¹ Some of these concerns were likely overblown, but they illustrate that domestic interest groups perceived backsliding as a real danger. In the area of parental leave, fears of benefits cuts proved justified.

The above examples obviously do not constitute a rigorous empirical test of our theory. Instead, they help explain the underlying mechanism of human rights backsliding. Previously mentioned patterns that have puzzled observers might be added to these illustrations: rights omitted from the UDHR becoming less popular in national constitutions than they had been before;¹⁰² and rich democracies sometimes worsening their behavior following treaty ratification.¹⁰³ In the next Section, we present a simple formal model to demonstrate that backsliding is theoretically coherent and plausible.

III.

THE MODEL

We argue that existing explanations of how human rights norms are transmitted to domestic policy predict both an improvement in the practices of low-performing states and a reduction in human rights protections in high-performing states. The model that follows allows us to make this point in a formal way and makes transparent the assumptions embedded in our analysis.

For purposes of illustration, we focus on pretrial detention, but the spirit of our argument applies to any human rights issue, including economic and social rights as well as civil and political rights. The magnitude of the effects we describe depends on a variety of factors, including the issue area at hand, the nature of the international norm, and the domestic politics of the receiving country, among others. For clarity of exposition, we first present a simplified model in this Section, and discuss these and other extensions and variations in the next Section.

101. For an overview of these concerns, see EHRENKRONA, *supra* note 90, at 45–51; *see also* Barbara Hobson, *Kön och misgynande: Svensk jämställdhetspolitik speglad i EG-domstolens policy, in* LJUSNANDE FRAMTID ELLER ETT LÅNGT FARVÅL? DEN SVENSKA VÄLFÄRDSSTATEN I JÄMFÖRANDE BELYSNING 174, 175–76 (Agneta Stark ed., 1997) (“Sweden is the archetype for the social democratic welfare model, and appears as the country that, because of the generous integration policies of the EU will lose most of its generous benefits and social services if the country is pressured—directly or indirectly—to adopt the minimum standard prevalent in countries with low taxes and social welfare costs”) (translated from Swedish by Helena Jung). For similar concerns, see also *id.* at 211–12. For other interesting examples of backsliding and backlash arguments, see CLIFFORD BOB, *THE GLOBAL RIGHT WING AND THE CLASH OF WORLD POLITICS* (2012).

102. *See* Elkins, Ginsburg & Simmons, *supra* note 10, at 81.

103. *See* SIMMONS, *supra* note 1, at 281–82 (suggesting that among high rule-of-law states, countries that have ratified the Convention Against Torture report a higher prevalence of torture than countries that have not ratified the Convention).

We begin by imagining how states would adopt a human rights policy in the absence of any international legal obligation. As with any policy, domestic decision makers must weigh the benefits and costs of their choice. We focus on the payoffs to the decision makers and assume that these payoffs include both private costs and benefits, and costs and benefits for society as a whole.¹⁰⁴ Costs and benefits to society as a whole are included in this calculation because in modern democracies, and in many autocratic states as well, leaders who pursue unpopular policies quickly lose power.¹⁰⁵

The tradeoffs associated with pretrial detention are familiar. Lengthy detention fundamentally limits an individual's liberty. When detention is imposed not only before a criminal conviction, but before a trial has even begun, it affronts fundamental notions of fairness such as the presumption of innocence. Moreover, detention imposes significant material costs on society as a whole, as it is expensive to incarcerate people for long periods of time. On the other hand, pretrial detention gives the state time to collect evidence needed for an effective investigation. In addition, it prevents individuals suspected of criminal behavior from fleeing. Pretrial detention is widespread—25 percent of people kept in European prisons are in pretrial detention, and the figures are much higher in other parts of the world.¹⁰⁶

Political leaders may want to limit pretrial detention for several reasons. First, limited pretrial detention reassures citizens that they (and members of their families, their political movements, ethnic groups, and communities more generally) are fundamentally free, and will not be imprisoned for long periods without a trial. Second, it frees up some state resources, allowing politicians to move funds from prisons to schools, social benefits, and other popular programs. Finally, government leaders may support limiting pretrial detention out of a sense of principle, because they believe that everyone is entitled to a speedy trial.

104. This assumption does not limit the applicability of our analysis. We require only that policy decisions are influenced by public perceptions and preferences. Backsliding would also result from a model involving only societal costs and benefits, but we believe it is more realistic to assume that political leaders make decisions in part because of private costs and benefits.

105. See Linos, *supra* note 1; LINOS, *supra* note 1; see also Weeks, *supra* note 42.

106. Thomas Hammarberg, *Excessive Use of Pre-Trial Detention Runs Against Human Rights*, THE COUNCIL OF EUROPE COMMISSIONER'S HUMAN RIGHTS COMMENT (Aug. 18, 2011, 9:12 AM), http://commissioner.cws.coe.int/tiki-view_blog_post.php?postId=169. The percentage of prisoners that are pretrial detainees by country is as follows: Liberia (97%), Mali (89%), Benin (80%), Haiti (78%), Niger (76%), Bolivia (74%), Congo-Brazzaville (70%). The United States has the world's highest number of pretrial detainees (about 476,000), and the fourth highest rate of pretrial detention (158 per 100,000). OPEN SOCIETY JUSTICE INITIATIVE, THE SOCIOECONOMIC IMPACT OF PRETRIAL DETENTION: A GLOBAL CAMPAIGN FOR PRETRIAL JUSTICE REPORT 16 (2011), available at http://www.unicef.org/ceecis/Socioeconomic_impact_pretrial_detention.pdf. By region, Asia has the highest percentage (47.8%), followed by Africa (35.2%). Mark Shaw, *Forward: Reducing the Excessive Use of Pretrial Detention*, JUSTICE INITIATIVES: PRETRIAL DETENTION, Spring 2008, at 13, available at <http://www.opensocietyfoundations.org/publications/justice-initiatives-pretrial-detention>.

On the other hand, political leaders may find that there are costs associated with limiting pretrial detention, especially from the perspective of the government currently in power. Limits on pretrial detention can force a government to release persons it suspects of criminal activity before it has had a chance to build a proper case against them. Or, alternatively, the government may need to devote more resources to prosecution efforts to ensure that cases are ready on time. Authoritarian regimes may enjoy the ability to freely detain opponents in order to limit anti-government protests, stay in power longer, and extract more private benefits. This basic balancing of priorities takes place in every country.¹⁰⁷ For our purposes, the specifics are not critical. We need only assume that leaders benefit from providing some protection of the relevant human right, but that at some point the benefits of further protection are outweighed by the costs. This is true for virtually every right discussed at the international level, including economic and social rights.¹⁰⁸

We index each country's policy on pretrial detention with the variable S , where $S > 0$. A higher value of S represents a regime that values individual liberties more greatly, and thus limits pretrial detention more severely.

We define a function, B_i , that represents the political benefits of more restrictive pretrial detention policies to the government of country i . We assume B is a simple linear function of S :

$$B_i = 2S_i, S > 0$$

Intuitively, the more limited pretrial detention is, the more support the government enjoys from citizens who value their own civil liberties and the civil liberties of others residing in their country.

We also define a function, C_i , that represents the costs of more restrictive pretrial detention policies to the government of country i . Like B_i , C_i is increasing in S . However, we expect steep increases in the costs of limiting pretrial detention as a government reduces pretrial detention from years to months to days to hours. If pretrial detention were extremely limited, or abolished altogether, the government would need to devote enormous resources

107. See generally MALCOLM M. FEELEY, COURT REFORM ON TRIAL: WHY SIMPLE SOLUTIONS FAIL 156–88, 239–42 (2013); R.A. Duff, *Pre-Trial Detention and the Presumption of Innocence* (University of Minnesota Law Sch. Legal Studies Research Paper Series, Working Paper No. 12-31, 2012).

108. The rhetoric surrounding rights (including speech rights) is often absolutist, resisting the notion that a particular standard is the result of competing priorities. This is an incorrect description of the vast majority of rights. The rights of criminal defendants, for example, are inevitably balanced against a societal interest in public safety; the rights of citizens to participate in the democratic process are balanced against concerns about the integrity of the voting system and administrability concerns; the rights of minority groups to be free from discrimination are balanced against a desire to avoid providing those same groups with unjustified advantages. Another way to make the same point is to observe that a human rights issue comes to be of interest only when there is some disagreement about the appropriate standard. There can only be disagreement when there are competing priorities. See Part IV.B.1 for a discussion of human rights standards that are implemented as “absolutist” norms.

to conduct very speedy criminal investigations, and would still be unable to detain most criminals. Thus, we posit that costs increase quadratically with S :

$$C_i = \lambda_i S_i^2$$

where $\lambda_i > 0$ reflects the idiosyncratic cost of human rights to country i .

A higher λ_i signals that the country's leaders view human rights as more costly to provide. This may be, for example, because they have a strong preference for crime control that limited pretrial detention will make more difficult, because they believe extensive pretrial detention will allow them to remain in power longer, because of their country's unique history, or for some other reason.

The λ_i variable, then, captures the different preferences among countries and leaders with respect to human rights. Simply put, as λ_i increases, the cost of any particular pretrial detention regime also increases, making it less attractive.

Our results would be largely the same with many other functional forms for both costs and benefits. The key is that there be some balancing of costs and benefits that leads to an interior solution with neither an infinite provision of the right nor a complete denial of it. We could have, for example, a convex cost curve (as we do) and a concave benefits curve (rather than our present linear one). This would complicate the arithmetic, but would not change the fundamental results. Indeed, the qualitative results would remain the same for any set of continuous, monotonic curves where curve B crosses curve C only once, from above.

In the absence of international law, each country chooses S_i , the level of freedom from detention without trial it will provide, so as to maximize its resulting gains. We define a leader's welfare function, W_i , as the difference between the benefits and costs of any particular S :

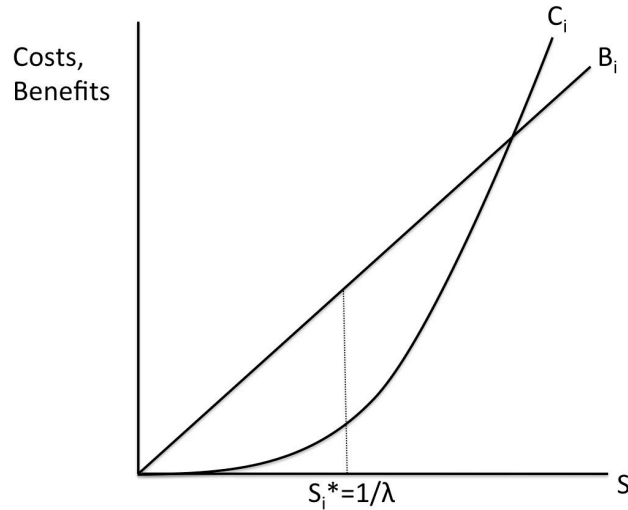
$$\begin{aligned} W_i &= B_i - C_i \\ &= 2S_i - \lambda_i S_i^2 \end{aligned}$$

Differentiating with respect to S to identify the first-order conditions and simplifying yields the optimal pretrial detention level:

$$S_i^* = 1/\lambda$$

This result, as one would expect, indicates that as λ increases, the level of human rights chosen by the state decreases. It is illustrated in Figure 1, with the optimal policy, S_i^* located at the point where the costs and benefits curves are furthest apart. This optimal point would be different for different countries.

FIGURE 1: Domestic Human Rights Policy in the Absence of International Law



Now consider how an international human rights agreement might impact this outcome. When a relevant international human rights norm, k , emerges, how is the domestic policy decision impacted?¹⁰⁹ We need to consider two cases. In the first, the international norm is higher than the equilibrium, S^* , that a country would otherwise choose. The second case is the opposite situation, where the international norm is lower than the domestic choice in the absence of an international influence.

To continue with our example, let us assume that an international standard emerges that countries may detain people without trial for a maximum of twenty months. Currently, while a right to a trial without undue delay is enshrined in many international agreements and in international custom, there is no specific maximum on pretrial detention periods to be applied across-the-board.¹¹⁰ And there is significant variation in national practices, even among

109. We put aside the important question of what it takes for such a norm to exist. It might require a treaty to which the state is a party, or perhaps it is enough that the treaty be widely signed, even if the state in question does not join. It could be that no formal treaty is required—a norm promulgated by an international organization might be sufficient.

110. See *Practice Relating to Rule 100. Fair Trial Guarantees*, INT'L COMM. OF THE RED CROSS, http://www.icrc.org/customary-ihl/eng/docs/v2_rul_rule100_sectionf (last visited Feb. 26, 2014); OFFICE OF THE HIGH COMM'R FOR HUMAN RIGHTS, HUMAN RIGHTS IN THE ADMINISTRATION OF JUSTICE: A MANUAL ON HUMAN RIGHTS FOR JUDGES, PROSECUTORS AND LAWYERS 267–71 (2003), available at <http://www.ohchr.org/Documents/Publications/training9Titleen.pdf>. It is possible to read the case law of the Human Rights Committee to support the position that countries may only detain people for up to twenty months, but we are greatly simplifying matters for clarity of exposition. The Human Rights Committee has held that a delay of twenty-four months between arrest and trial violates article 14(3)(c) of the ICCPR. The Committee has reached the same

rich democracies in Europe. For example, the United Kingdom limits pretrial detention to six months, and as of 2009, the average length of pretrial detention was thirteen weeks.¹¹¹ In contrast, France limits pretrial detention to four years, and the average length of pretrial detention in 2005 was almost nine months.¹¹² If an international norm limiting pretrial detention to twenty months emerged, France would offer lower protections than this international norm, while the United Kingdom would offer higher protections.

A. International Norm Higher than Domestic Equilibrium ($k > S^$)*

Consider first the case in which the international norm demands a higher level of human rights protection—more limited pretrial detention in our illustration. As already mentioned, we assume this norm matters because it bolsters the political power of interest groups and politicians within the country seeking an increase in human rights. In our case, that means that interest groups and politicians who support more limited pretrial detention gain support. Interest groups that favor civil liberties generally may decide to focus their campaigns on pretrial detention, because the international standard provides them with an added argument. And politicians who support limited pretrial detention will be able to point to the international standard come election time, and argue that their policy position is mainstream and sensible.

We model this as a discontinuity in the benefits curve facing the state. If the international threshold is satisfied, political leaders gain a discrete increase in support, which we call D .

There are several other ways to model this benefit that would give us similar results—for example, we could make the slope of the benefits curve steeper. We think modeling the gain as a discrete increase makes our model more realistic. Politicians who reach the international standard can more convincingly argue that their policies are sensible ones, and may gain external validation from international organizations and foreign governments. In contrast, politicians who merely move towards the international standard, but do not actually reach it, will have a harder time using the international standard as an external benchmark.

We represent the effect of an international norm with a modified benefits curve, B_i' :

$$\begin{aligned} B_i' &= 2S_i, \text{ if } S < k \\ &= 2S_i + D, \text{ if } S \geq k \end{aligned}$$

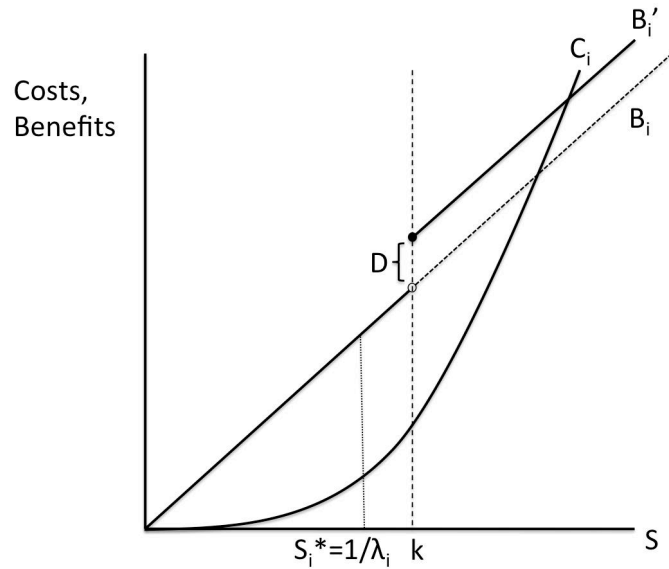
result for delays of twenty-nine months, thirty-one months, and longer periods of time. However, it has held in the case of Kelly that a delay of eighteen months was not unreasonable. *See id.* at 267–69.

111. FAIR TRIALS INT'L, DETAINED WITHOUT TRIAL: FAIR TRIALS INTERNATIONAL'S RESPONSE TO THE EUROPEAN COMMISSION'S GREEN PAPER ON DETENTION 49 (2011), available at <http://www.fairtrials.org/documents/DetentionWithoutTrialFullReport.pdf>.

112. *Id.* at 52.

The cost and benefits curves then look as follows:

FIGURE 2: Human Rights Behavior with International Law



Notice that the international norm is to the right of the state's ex ante preferred position. This reflects the fact that the international norm is more protective of detainees than is S_i^* . The arrival of this norm presents the state with a new possibility. By moving from S_i^* to k , state leaders incur greater costs, but also benefit from the discrete increase in benefits that occurs at k .

Because S_i^* is the best outcome for state leaders in the absence of international law, and because the international norm has not affected their payoff anywhere to the left of k , we know that these leaders prefer to remain at S_i^* rather than move to any point between S_i^* and k . Similarly, because the slope of C_i increases as we move to the right, we know that choosing a position to the right of k , $k + \epsilon$, for any $\epsilon > 0$ is worse for the state than choosing k .¹¹³

The state, then, will maximize its welfare, W_i , by choosing either S_i^* or k . The relevant welfare gains are represented as follows:

$$\begin{aligned} W_i(S_i^*) &= 2S_i - \lambda S_i^2 \\ &= 1/\lambda \end{aligned}$$

$$\begin{aligned} W_i(k) &= 2S_i + D - \lambda_i S_i^2, \text{ where } S_i = k \\ &= 2k + D - \lambda_i k^2 \end{aligned}$$

113. We know this because the slope of C at the pre-international law optimum, S^* , is equal to the slope of B (this is necessarily true if S^* is the optimum). Moving to the right from there, the slope of C increases while the slope of B remains constant and the slope of B' is the same as that of B . It follows that the net benefits decrease as we move to the right.

Comparing these two outcomes, the state will choose the larger one. In other words, if:

$$1/\lambda_i > 2k + D - \lambda_i k^2$$

then the state remains at S_i^* and is unaffected by international law. If the opposite is true, the human rights norm is effective and the state moves to k . We are interested, therefore, in the relative values of k and λ at which the state is indifferent between improving its conduct and remaining at S_i^* .

The critical value of k at which a state with costs λ_i is indifferent is given by solving:

$$\begin{aligned} 1/\lambda_i &= 2k + D - \lambda_i k^2 \\ \lambda_i k^2 - 2k + (1/\lambda_i - D) &= 0 \end{aligned}$$

Solving this for k yields:¹¹⁴

$$\bar{k} = 1/\lambda_i + D^{1/2}/\lambda_i^{3/2}$$

This relationship between k and λ_i identifies how states will react to the international norm. Specifically, a state with cost λ_i will respond to the norm by improving its human rights conduct (moving to k) as long as the chosen k is less than \bar{k} :

$$k < 1/\lambda_i + D^{1/2}/\lambda_i^{3/2}$$

If, on the other hand, the international norm is greater than \bar{k} , it will have no effect on the country. This means that a human rights norm that is too ambitious will leave states with poor human rights records unaffected.¹¹⁵

For example, article 5(a) of CEDAW, which requires that “State Parties shall take all appropriate measures . . . to modify the social and cultural patterns of conduct of men and women, with a view to achieving the elimination of prejudices and customary and all other practices which are based on the idea of the inferiority or the superiority of either of the sexes or on stereotyped roles for men and women,”¹¹⁶ is unlikely to have much impact on states with the worst records on women’s rights.

This relationship between the content of the norm and the set of countries impacted suggests that there is some reason to produce treaties that are less ambitious in the hope of more effectively influencing states with poor human rights records.¹¹⁷ For example, many European governments have criticized the United States for watering down human rights standards in the ongoing negotiations over a potential multilateral treaty that would regulate trade in

114. To arrive at this result we use the quadratic formula to solve $ax^2 + bx + c$: $(-b \pm (b^2 - 4ac)^{1/2})/2a$. This yields (omitting subscripts) $k = 1/\lambda \pm D^{1/2}/\lambda^{3/2}$. We know, however, that $k > S_i^*$ and $S_i^* = 1/\lambda_i$. Therefore, $k > 1/\lambda_i$ and so we must add the second term in the expression for k rather than subtract it.

115. See *infra* Part V.B.3.

116. Convention on the Elimination of All Forms of Discrimination Against Women, *supra* note 25, at art. 5(a).

117. An important caveat to this point is developed below, where we explain that states may wish to include high standards to prevent backsliding. See *infra* Part V.B.3.

conventional weapons.¹¹⁸ However, our theory suggests that lower standards may in fact have a chance to generate an improvement in human rights that would not be possible with a tougher treaty.¹¹⁹

Our model also suggests that if we are able to reduce the costs of adopting better human rights, more states will be responsive to an international norm. In particular, if the cost of better human rights conduct is reduced (i.e., smaller λ) then more states will move to satisfy the norm and states further away from the norm will do so as well.

Similarly, an increase in the benefits received by states that satisfy the norm (D) increases the range of states that are responsive and does so by making states farther away from the norm move to k .

Notice that these results accord well with our intuitions. An increase in D makes more states responsive to any particular choice of k ,¹²⁰ and an increase in the cost of granting human rights (i.e., higher λ) makes it less likely that a state will respond to an international norm.

B. International Norm Lower than Domestic Equilibrium ($k < S^$)*

The discussion above considers the case in which the chosen international norm is more protective of human rights than the national policy in the absence of international law. We now turn to consider the alternative situation in which the international norm is less protective than the ex ante domestic policy. In most human rights discussions, it is implicitly assumed that an international norm, even if it is effective, will have no impact on states that provide a higher level of human rights protections. In what follows, we first consider the assumptions necessary to generate this lack of impact, and then turn to show how a change in those assumptions might lead to a different outcome.

Figure 3 illustrates the case in which the state's preferred position in the absence of international law, S_i^* , is more protective of detainees' rights than the international norm, k . The presence of the norm increases the benefits received by the state's leaders because they enjoy the discrete gain, D , associated with meeting the norm. It does not, however, change a state's human rights conduct. Graphically, this is evident because the point at which the state maximizes its gains is the point at which the slope of B_i (or B_i'), which reflects the marginal benefits from increased human rights protection, is equal to the slope of C_i , which is the marginal cost of increased protections. The

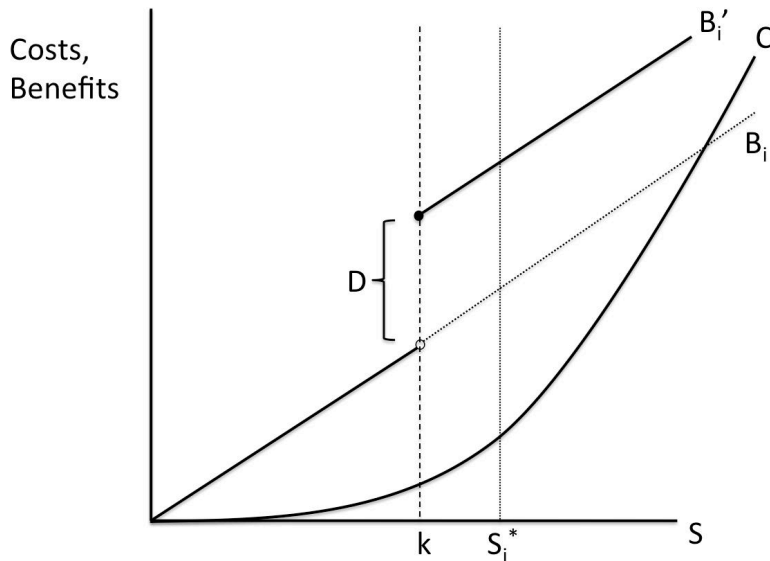
118. See Nick Hopkins, *UK Presses US on Human Rights Clause in Arms Trade Treaty*, THE GUARDIAN (July 1, 2012), <http://www.guardian.co.uk/world/2012/jul/02/human-rights-arms-trade-treaty>.

119. It is sometimes observed that treaties must at times be weakened in order to get more states to join the treaty regime. Our point is somewhat different. Making a treaty weaker may alter the subset of participating states whose behavior is affected. Specifically, weakening a regime may influence those with poor human rights records more than could be done with a tougher treaty.

120. Notice also that if $D = 0$, $k = 1/\lambda = S^*$, meaning that we are back in the base case without international law. This would also be the case if international law had no effect.

discontinuity in the B_i curve at the point k does not affect the slope of either the B_i or the C_i curve, so the optimum is unchanged.

FIGURE 3: Current View of Human Rights Behavior with International Law



This representation of a human rights norm is consistent with the dominant view that establishing a norm will help to improve the performance of some states that previously fell short of the norm, but will have no impact on those that exceed the norm at the time of adoption. To continue with our illustration, the conventional wisdom is that the United Kingdom should not change its laws limiting pretrial detention to six months if an international norm emerges that sets maximum detention to twenty months.

Given our theory of human rights policy formation, however, there is reason to wonder if this result is correct. Our theory assumes that an international norm affects domestic politics by providing support to domestic groups that favor an outcome similar to that specified in the international norm. However, the focal point created by the international norm could also exert some influence in states with human rights practices above the international standard. The increase in political influence enjoyed by groups that support a domestic policy similar to the international norm must come at the expense of other groups, possibly including groups that prefer a higher level of human rights.

This point is illustrated in the examples given in Part II. Consider how it would come about in the context of pretrial detention. For countries with limited protections for detainees, our theory suggests that an international norm

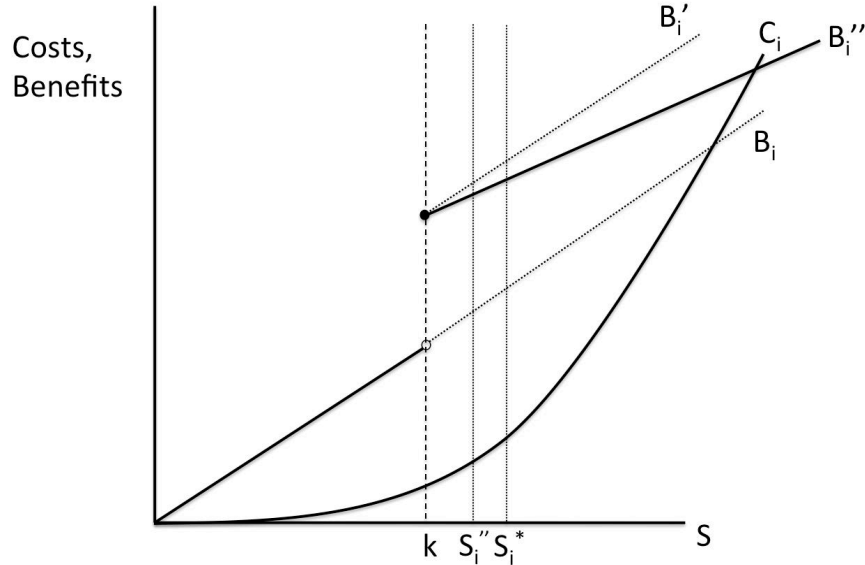
specifying some required (or perhaps only recommended) level of protection for detainees can increase the influence of local interest groups that seek domestic policies with similar protections. This increase in influence can prompt political leaders to make different policy choices. We model this political dynamic as a discrete gain to countries that satisfy the international norm.

Now consider a state with existing protection in excess of that required by the international norm. To continue with our example, assume that before the international standard emerges, the British government is persuaded that its current policy of detaining individuals for a maximum of six months without trial represents the right balance between the liberties of the individual detainee and the ability of the state to properly investigate crimes. If an international standard emerges setting maximum detention at twenty months, the United Kingdom would face no legal obligation to change its policies. But the logic we propose here is political. The United Kingdom, like all other countries, has interest groups engaged in a debate about the merits of increased protection for detainees. When an international norm emerges, groups that favor stricter crime-control policies may gain influence. These groups could point to the international norm to support their claim that the appropriate balance is less protective of detainees than the status quo. In contrast, advocacy groups supportive of extensive civil liberties protections and limited pretrial detention might start to believe that the British government is doing relatively well on pretrial detention issues. They might start devoting their efforts to areas where the United Kingdom is performing below the international standard. Similarly, citizens who once favored extensive protections of civil liberties and limited pretrial detention might start doubting their earlier commitments, and rewarding their government less for pursuing liberal policies far outside the global mainstream.

Figure 4 shows how this backsliding could happen. Once the international norm, k , is established, the perceived benefits of marginal movements to the right of k are reduced. That is, the international norm signals that the balance between the human right in question (e.g., civil liberties) and other concerns (e.g., crime control) recommends a policy located at k . That information filters through the domestic political process in a way that reduces the marginal benefits enjoyed by political leaders when they adopt policies to the right of k .¹²¹ We represent this change in benefits as a change in the slope of the B_i' curve to the right of k , shown by the line B_i'' .

121. A similar result could be achieved by modeling the change as an increase in costs.

FIGURE 4: Human Rights Behavior with International Law and Backsliding



The state reacts to this new benefits curve by adjusting its provision of human rights so as to maximize the net gains from its policy. This entails a move to the left, to S_i'' .

The same result can be shown more formally:

$$\begin{aligned} C_i &= \lambda_i S_i^2 \\ B_i &= 2S_i, \text{ if } S_i < k \\ &= 2k + D + (S_i - k), \text{ if } S_i \geq k \end{aligned}$$

The benefits curve includes $2k + D$, the benefit enjoyed by the state if it adopts policy k , plus the additional benefits of moving to the right beyond k . Note that, as drawn, the slope of the B_i'' curve is one, less than the slope of the B_i curve, which is two. This captures the notion that the international norm may provide a kind of anchor below that country's status quo ante.

Any country that chooses a level of human rights less than k in the absence of international law will be affected in the same way as already discussed in Part III.A. That is, states located to the left of k , but sufficiently close to it, will migrate to k . Those located farther to the left, i.e. states with particularly poor human rights records, will not be affected.

Because the slope of B_i'' has changed, there is now an effect felt by states to the right of k . In the absence of international law, these states choose $S_i^* = 1/\lambda$, as shown in Figure 4.

With the international norm in place, they maximize the difference between benefits and costs:

$$W_{Li} = 2k + D + S_i - k - \lambda S_i^2$$

Differentiating with respect to S allows us to identify the first-order conditions for this problem and yields:

$$S^*_{ILi} = 1/2\lambda$$

Notice that this is to the left of the prior outcome of $S_i^* = 1/\lambda$.

S^*_{ILi} may be the optimal policy for the state's leaders, but we must consider one additional constraint. Given that states started off to the right of k , we know that they will not respond to the international norm by moving to the left of k . Doing so would sacrifice the benefit D and place them on the original B_i curve. We know this is inferior to an outcome in which the state chooses to locate itself at k .

States will, however, move to k if the welfare from being at k exceeds that of being at S^*_{ILi} . Yet, from the first order conditions we know that if $1/2\lambda_i$ is larger than k , the state is better off at $1/2\lambda_i$. If the opposite is true, and $k > 1/2\lambda_i$, then the state will always be better off at k , which we know because prior to the arrival of international law the state preferred $S^* > k$, which implies that moving to the left from point k reduces the payoff to the state.

All of this means that the state will move to the larger of $1/2\lambda_i$ and k . In either case we get backsliding. To summarize, it is the flattening of the B curve that is doing the work in our backsliding theory. A completely flat B curve would lead every country that would otherwise provide higher protection than the international standard to backslide all the way back to the international standard.¹²²

Comparing the two cases, it is worth noting that the positive effects of the international norm are felt only by states that are below but close to the norm while the backsliding effects are felt by all states that are above the norm. This does not necessarily mean that the norm does more harm than good, of course, because we would need additional assumptions in order to make a prediction about the number of countries in each category and the magnitude of the movements, not to mention how to aggregate outcomes across states.

IV.

EXTENDING THE THEORY OF BACKSLIDING

For clarity of exposition, we intentionally kept the backsliding model presented above simple. In this Section, we offer some further analysis of the

122. This extreme case of a B' curve that is flat to the right of k is worth considering, not because it is realistic, but because it can clarify this logic. A B' curve that is flat to the right of k reflects a norm strong enough to eliminate any governmental benefits from moving farther to the right. It is difficult to imagine such a norm in practice, but it might be one that is not merely a floor, but also a ceiling, meaning it specifies a precise level of treatment that is required. So, for example, if (contrary to fact) the international norm specified that states must limit pretrial detention to exactly twenty months, and this norm were very powerful, the B' curve might be flat. If the international norm causes the B curve to become flat to the right of k , then every country that would otherwise provide a higher level of protection will backslide all the way to k .

theory and consider several extensions. We first outline how domestic factors, such as democracy and the role of the judiciary, influence the impact of backsliding. We then examine how changes to the form and content of a norm might matter. We consider the following: the use of minimum standards (floors), maximum standards (ceilings), and prohibitions; the legal form of the norm (hard law versus soft law); norms that identify multiple standards; and vague norms and uncertainty with respect to national compliance. Finally, we discuss how alternative mechanisms that explain how international norms influence state behavior could also lead to backsliding.

A. Domestic Factors Influencing Backsliding

A broad implication of our theory is that once an international standard comes into place, state behavior should tend toward this standard as poor performers are pulled upward and top performers are pulled down. It does not predict, however, that states will fully converge on the international norm or that all similarly situated countries will necessarily move in the same fashion. Instead, countries differ in their receptiveness to international norms.

We argue here that factors that make some countries especially receptive to international norms, and thus especially likely to improve their behavior when an international norm is set at a high level, also contribute to a high risk of backsliding. Countries that are highly integrated into the international system and whose governments regularly participate in international meetings and take on international obligations, are more likely to be influenced by international norms than countries that are more isolated from the international system, both positively and negatively.¹²³ For example, because the United States resists international norms more strongly than many other rich democracies do, the risk of backsliding is reduced, as is the potential for an upward pull.¹²⁴ We also expect democratic countries to be more receptive to international norms than authoritarian ones, because the domestic politics mechanism we just described relies on pressures from voters, interest groups, and other key constituencies.¹²⁵ The heightened receptivity of democracies to international norms should influence countries both when the international

123. Other causes of rights stagnation and diminutions include changing views of the impacts of rights, economic and political shocks, treaty fatigue, and a desire to avoid creating new international institutions. *See, e.g.*, Benedikt Goderis & Mila Versteeg, *Human Rights Violations After 9/11 and the Role of Constitutional Constraints*, 41 J. LEGAL STUD. 131 (2012) (The authors argue that after 9/11, the human rights records of the United States and its allies deteriorated due to a general consensus among democracies to trade some measure of human rights protection for extra security. This deterioration cannot simply be explained by the amount of threat terrorist groups posed to each country; it is better explained as a (perhaps momentary) shift in understanding about the relative value of human rights and national security.).

124. *See* LINOS, *supra* note 1, at 36–38 (explaining that Americans are less supportive of a variety of international efforts than citizens of other countries).

125. *See id.* at 18–26; Linos, *supra* note 1, at 679.

standard is set high relative to their current level of domestic protections, and when it is set relatively low.

It follows that democracies that are highly integrated in international and regional systems face the greatest risks of backsliding. For example, countries like the United Kingdom and Sweden, which are highly integrated in European regional systems, face high risks of backsliding on issues where Anglo-Saxon and Scandinavian legal traditions offer very different (and more robust) protections from those prevalent among continental European states. Similarly, strong regional human rights ties within the Americas can place top performers in this region at risk of backsliding.

We note one curious feature of this prediction. If one assumes that democracies generally have better human rights records than autocracies, the asymmetry between the responses of these two groups to a norm may at times frustrate human rights efforts. A norm aimed at improving the conduct of low performers (by assumption, predominantly autocracies) will have a relatively small positive influence on those same low performers but will have a relatively large backsliding impact on high-performing democracies. One can debate whether this compromise—modest upward movement among low performers and more significant downward movement among high performers—is desirable. Moreover, other mechanisms may also be at work, mechanisms that operate asymmetrically and are designed to work only in the direction of expanding rights protections. But, at a minimum, the risk of backsliding should give one pause before advocating for such a norm.

The above discussion identifies democracies as susceptible to backsliding because we expect democracies to be more responsive to domestic interest groups and, therefore, more responsive to international norms. This point can be generalized. Any feature of a domestic political system that affects the force with which international norms are transmitted to domestic policy can impact the level of backsliding.

For example, backsliding may be less of a concern where rights are largely determined by the judiciary rather than by the legislature or the executive. In our formal model, we use pretrial detention as an illustration. In the United States, the judiciary plays a key role in determining the scope of pretrial detention. Where this is true—where the content of a right is determined by courts rather than other actors—domestic policy is subject to fewer political influences.¹²⁶ It follows that—with the possible exception of legally binding norms that domestic courts are required to respect and enforce—international human rights norms will have less influence on domestic policy, and consequently pose less of a risk of backsliding.¹²⁷

126. See LINOS, *supra* note 1, at 27–29; Linos, *supra* note 1.

127. However, especially in Europe, there is a growing trend of national judges interpreting domestic constitutional rights consistently with international human rights standards. As a result, a

The key point here is that the risk of backsliding goes hand-in-hand with the potential for international norms to improve outcomes. The politics that lead to backsliding are fundamentally the same as the politics that allow international norms to enhance domestic policies. Any influence that strengthens transmission from international norms to domestic policies will promote human rights policies in some countries and increase the risk of backsliding in others.

B. How the Content and Form of the Norm Matters

1. Floors, Ceilings, and Prohibitions

Our discussion so far has assumed that the norms at issue are floors and not ceilings. That is, we have assumed that the relevant international norm requires states to meet some minimum threshold but allows them to exceed it if they wish. This is a fair characterization of most key international human rights obligations.¹²⁸ The ICCPR provides a right to freedom of speech, for example, but countries are at liberty to provide more expansive rights, as the United States and many others do.¹²⁹ Similarly, the Covenant establishes a right for men and women of marriageable age to marry, but this does not prohibit states from allowing same-sex couples to marry as well.¹³⁰

One consequence of our theory of norm transmission from the international to the domestic context is that backsliding can take place even though no state is obligated to reduce its level of human rights protections. We acknowledge that when international norms require (even in a non-binding fashion) a floor, the upward pull of those norms may be stronger than the downward pull we label backsliding. The magnitude of the difference depends on the particulars of how the norm affects domestic decision making. To the extent one believes that the formal legal commitment to meet a specified floor is what causes a change in behavior, backsliding is a relatively small problem. But if one believes (as we do) that the signal provided by the international norm can also provide a political cover for those opposed to a particular expansion of rights, backsliding should be taken seriously.¹³¹

A closely related issue is whether it is possible to minimize backsliding through careful drafting. Our sense is that while it may be possible to mitigate the backsliding effect, alternative forms of drafting that do not change the substantive content of the norm are unlikely to make much of a difference.

national court may serve as an additional mechanism for backsliding rather than acting as a buffer against it. See initial case studies in this Article, *supra* Part II.A–C.

128. See *supra* notes 25–30.

129. See International Covenant on Civil and Political Rights, *supra* note 26, at art. 19.

130. See *id.* at art. 23.

131. As shown in Part III, Figure 4, the pull of the minimum requirement corresponds to the upward shift in the *B* curve while the empowerment of opposition to enhanced rights corresponds to the flattening of the slope of the same curve.

Instead, as we outline in Part V.A, bigger changes in treaty design and international organization membership will be needed. Because our transmission mechanism relies on a political rather than a legal dynamic, technical and legal changes to a treaty or other instrument are not likely to have a large impact on how the content of that instrument is used in domestic political debates. A drafter could, for example, make explicit that a treaty's requirements constitute a floor and that states are free to exceed that floor. Indeed, this is often done, as noted above. We are not persuaded, however, that this has much impact on the ways the treaty is used by opponents of the right at issue.¹³² Opponents will still be able to point to the treaty as evidence that the international community does not believe that a particular behavior is required.¹³³

One could also write an international norm as an exact target, rather than as a floor. An international agreement might recommend, for example, that states provide exactly six years of public education, rather than recommending that states provide a minimum of six years of public education. Such a norm would be both a floor and a ceiling, and would presumably pull states in both directions. This is not an interesting example of what we call backsliding. Backsliding is theoretically interesting and normatively troubling precisely because most human rights norms are not designed as exact standards, but as minimum floors.

Some human rights norms are better characterized as absolute prohibitions rather than a floor or a ceiling. The prohibition on torture, for example, might fit this description, as might the prohibition on genocide.¹³⁴ In both cases, the international norm imposes a complete ban on the activity. At first glance, it may seem that these norms are not subject to backsliding because there is nothing a country could do to provide a higher level of protection than that required by a complete prohibition.

If, however, one views torture (or genocide) as simply a label assigned to that which is prohibited, the issue is more complicated, because the omission of a particular practice from the list of actions constituting torture opens the door to backsliding. Although there is no way for a country to adopt a stance toward torture that is stricter than the complete ban imposed by the Convention

132. Other mechanisms, such as the creation of optional protocols, the delegation of treaty interpretation to courts or other bodies that can increase standards over time, and the creation of regional agreements can all help reduce the risk of backsliding by creating multiple standards. We elaborate on these points in Parts IV.B.3 and V.A below.

133. Our view here reflects the political influence mechanism adopted throughout the paper. Alternative assumptions about how international norms influence domestic actions might lead to a different view of whether drafting is a promising tool for minimizing backsliding.

134. See Convention on the Prevention and Punishment of the Crime of Genocide, Dec. 9, 1948, 78 U.N.T.S. 277; see also David Luban, *Calling Genocide by Its Rightful Name: Lemkin's Word, Darfur, and the UN Report*, 7 CHI. J. INT'L L. 303, 305–06 (2006) (arguing that the overly legalistic reading of genocide by international courts is entirely divorced from the way most people think of genocide as the "crime of crimes" deserving absolute prohibition).

Against Torture or customary international law, it is possible for a country to define torture more broadly than these sources or to prohibit actions that go beyond the formal definition of torture.¹³⁵ For example, while the United Nations Convention Against Torture applies only to torture by or with the consent or acquiescence of a public official,¹³⁶ the Inter-American Convention to Prevent and Punish Torture proscribes torture by a broader set of actors.¹³⁷

If the international norm fails to prohibit, for example, a particular interrogation technique, domestic actors arguing in favor of using that technique may be able to point to international standards in support of their claim. They may be able to point to the international standard to argue that the appropriate trade-off between humane treatment of detainees and the need to carry out effective interrogations leaves room for the method in question. Indeed, when the United States was debating this exact topic during the first decade of the twenty-first century, both sides argued that their respective position was consistent with international law and norms.¹³⁸

Notice that this is not a point about the vagueness of the norm. It focuses instead on the boundaries of a prohibition, however clear they are. Genocide, for example, is defined in the Genocide Convention to include only actions targeted at “national, ethnical, racial or religious group[s].”¹³⁹ To the extent one believes that this convention has an upward pull on the behavior of states considering acts against any of the enumerated groups, it will also have a downward pull (backsliding) on states considering acts against groups that are not listed, such as groups targeted because of political ideology or wealth.

2. *Hard Law Versus Soft Law*

Our theory of backsliding rests on a particular view of the transmission mechanism through which international norms affect domestic policy. Specifically, international norms enhance the political power and influence of some domestic actors at the expense of others. This political transmission mechanism need not depend on the formal legal status of an international norm

135. See Oona A. Hathaway, Aileen Nowlan & Julia Spiegel, *Tortured Reasoning: The Intent to Torture Under International and Domestic Law*, 52 VA. J. INT'L L. 791, 798 (2012).

136. Convention Against Torture and Other Cruel, Inhuman or Degrading Treatment or Punishment art. 1, Dec. 10, 1984, 1465 U.N.T.S. 85.

137. Inter-American Convention to Prevent and Punish Torture art. 3, Dec. 9, 1985, O.A.S.T.S. No. 67 (entered into force Feb. 28, 1987); see also DIEGO RODRÍGUEZ-PINZÓN & CLAUDIA MARTÍN, THE PROHIBITION OF TORTURE AND ILL-TREATMENT IN THE INTER-AMERICAN HUMAN RIGHTS SYSTEM: A HANDBOOK FOR VICTIMS AND THEIR ADVOCATES 150–56 (Leonor Vilás Costa ed., 2006), available at http://www.omct.org/files/2006/11/3977/handbook2_full_eng.pdf.

138. See generally Jinee Lokaneeta, *Torture Debates in the post-9/11 United States: Law, Violence, and Governmentality*, 13 THEORY & EVENT (2010); see also Zachary R. Calo, *Torture, Necessity, and Supreme Emergency: Law and Morality at the End of Law*, 43 VAL. U. L. REV. 1591, 1594–97 (2009).

139. See Convention on the Prevention and Punishment of the Crime of Genocide, *supra* note 134, at art. II.

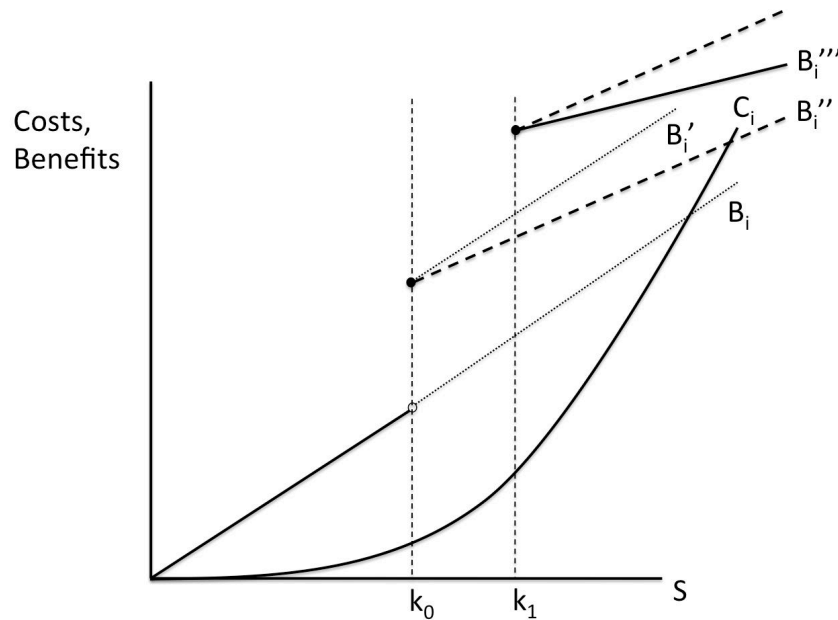
(i.e., whether the norm is binding or not), implying that both soft law and hard law norms can have an impact on policy and lead to backsliding. Among other consequences, this means that state consent is not necessary for a state to run the risk of backsliding.

To say that legal form is not determinative is not to say that it is irrelevant. The particular form that a norm takes may affect the influence the norm has over domestic actors, so formal legal status (among other things) may still be relevant. To the extent that formal international law exerts greater influence on the domestic sphere, it will have a larger impact than other norms that do not have this legal status.

The overall implication, then, is that the impact of norms and the risk of backsliding depend on the force with which international policies or commitments get translated into domestic political influence. The formal legal status of a norm may matter, as may the issue area, the source of the norm, the enthusiasm of the domestic political system for foreign models, and so on.

3. *Multiple Standards*

The analysis of backsliding is somewhat more complicated when multiple international norms are at play in a single issue area. The simplest case would feature several threshold levels of performance ranked from the easiest to the hardest to achieve. As we explain below, international treaties often have optional protocols that specify different performance thresholds. In our model, we represent the case of multiple norms with several discontinuities in the B curve, rather than just one. Instead of a single point, k , there would be several, labeled k_0 , k_1 , k_2 , and so on. A country that satisfied the first threshold, k_0 , would enjoy a discontinuous jump in benefits, just as in our base model. A country that also met the next threshold would see a further jump.

FIGURE 5: Human Rights Behavior with Multiple Standards

The results here are not so different from the base case of a single threshold. To the left of the first threshold, k_0 , and the right of the last one, k_1 , states will behave just as they do in our base model. Thus, states to the left of k_0 will improve their performance if they are already sufficiently close to the threshold. If they are not, the norm will have no impact on them. All states to the right of k_1 will backslide to some degree, as in the basic model.

The more interesting cases are states that are “in between” two thresholds, including states that were pulled up to a particular threshold from below. These states face an incentive to move up to the next higher threshold in order to benefit from the discrete jump in benefits. Some will also, however, experience some backsliding as a result of the lower threshold’s tendency to flatten the benefits curve.

The dynamic involved is best understood if we distinguish the two forces at work. First, the existence of a threshold to the left of where a country would otherwise be leads to backsliding, and that is what we observe if all thresholds are to the left of the country. The extent of the backsliding depends on the extent to which the slope of B is flattened. The more the curve flattens, the smaller the gain to a state that improves its human rights performance. Second, having a threshold to the right means that the state may wish to improve its performance enough to enjoy the discrete gain in benefits from meeting that threshold.

The state’s choice, then, is between moving to the higher threshold and backsliding in the direction of the lower one. Weighing these two options, the state must determine which is more favorable. The closer it is to the higher

threshold, the more likely it is to improve its performance. It is conceivable, then, that having two thresholds will cause some relatively high performers to improve while causing some relatively poor performers to backslide.

The optional protocol is one of several tools that help generate multiple standards in the human rights field. In many human rights areas, the main treaty sets a relatively low and uncontroversial standard, while one or more optional protocols articulate higher and more controversial standards. For example, the ICCPR forbids the death penalty specifically for juveniles and pregnant women,¹⁴⁰ while the Covenant's Second Optional Protocol abolishes the death penalty more generally.¹⁴¹

Another way to create multiple standards is to delegate the interpretation of key provisions to a treaty body or court that can alter a standard's meaning over time.¹⁴² For example, after the Human Rights Commission interpreted the prohibition on sex discrimination in Articles 2 and 26 of the ICCPR to include sexual orientation, it concluded that an Australian province's prohibition on consensual same-sex activity violated the ICCPR.¹⁴³ Still another way to generate multiple standards involves regional agreements that define fundamental rights differently from universal agreements. We examine this point in detail in Part V.B below.

4. *Vague Norms and Uncertainty*

Up to this point, the Article has assumed that both the international norm and the actions of states are observable. In practice, however, the international norm itself may be contested or uncertain, and policy implementation by a state may be difficult for others to observe. At one extreme, a norm with requirements so vague as to be meaningless will have no effect on state behavior. Similarly, if the conduct of the state implementing the norm is entirely unobservable, even by its own citizens, then the norm will not affect behavior.

Whether uncertainty stems from a vague norm or difficult-to-observe government policies, the effects on our model are the same. Rather than a discrete jump upward at k , the benefits curve has a more gradual increase in slope as it approaches k and then flattens once it passes to the right of k . This more continuous change in the curve reflects the fact that a state with policies

140. International Covenant on Civil and Political Rights, *supra* note 26, at art. 6(5).

141. Second Optional Protocol to the International Covenant on Civil and Political Rights, Aiming at the Abolition of the Death Penalty art. 1, Dec. 15, 1989, 1642 U.N.T.S. 414.

142. Andrew T. Guzman & Timothy L. Meyer, *International Soft Law*, 2 J. LEGAL ANALYSIS 171, 178 (2010).

143. See International Covenant on Civil and Political Rights, *supra* note 26, at arts. 17, 26; Human Rights Committee, *Toonen v. Australia*, *supra* note 67. *But see* Laurence R. Helfer, *Overlegalizing Human Rights: International Relations Theory and the Commonwealth Caribbean Backlash Against Human Rights Regimes*, 102 COLUM. L. REV. 1832, 1851–60 (2002) (detailing the increased likelihood of backlash as human rights expand).

close to that required by an international norm may be able to persuade some observers (including perhaps local citizens) that it is actually in compliance with the norm. Similarly, if the norm is vague, a state may be able to persuade some observers that its conduct is sufficient to satisfy the norm.

Because there is no discrete jump up in the benefits curve, some states may decide to improve their performance and yet not move all the way to k . As one would expect, vague norms or difficult-to-observe policies generate less upward pull than clear norms and easily observable domestic conduct. For high-performing states, there can be a similar reduction in the risk of backsliding. If domestic policy is difficult to observe, a state that is in compliance, but close to the international norm, may enjoy only a portion of the benefits of compliance. Graphically, this would be represented by a benefits curve that is still flattening to the immediate right of the international norm. This has the effect of reducing the risk of backsliding for states close to the international norm.

C. Other Theories of International Law's Influence

Several theories seek to explain how and why international law can induce governments to better protect human rights without coercive enforcement. Up to this point, we have focused on a mechanism we find particularly plausible. In our account, international norms shape the views of domestic interest groups, especially undecided voters.¹⁴⁴ These international norms can be used to support a domestic policy proposal, and can help persuade ordinary citizens that a given proposal is not a radical experiment but rather a mainstream, tried-and-true solution.¹⁴⁵ In this Section, we explain why other prominent theories of international law's influence might also predict backsliding.¹⁴⁶

One prominent theory of international law's influence, developed by scholars such as Professors Abram and Antonia Chayes and Professor Thomas Franck, states that treaties exert a "compliance pull" on states because states

144. See LINOS, *supra* note 1, at 36–66; Joseph M. Grieco et al., *Let's Get a Second Opinion: International Institutions and American Public Support for War*, 55 INT'L STUD. Q. 563 (2011); Michael Tomz, *Domestic Audience Costs in International Relations: An Experimental Approach*, 61 INT'L ORG. 821, 837 (2007); Linos, *supra* note 1; Geoffrey P.R. Wallace, *International Law and Public Attitudes Toward Torture: An Experimental Study*, 67 INT'L ORG. 105 (2013).

145. See *infra* Parts I, III.

146. A caveat: there are some scholars who reject the possibility that international law can positively affect state behavior under most circumstances. These scholars argue that strategic concerns and domestic pressures entirely unrelated to international law dominate leaders' calculations. See GOLDSMITH & POSNER, *supra* note 34, at 13. If this view is correct, and international law is too weak to influence state behavior for the better, then it is quite unlikely that international law could influence state behavior for the worse. We set aside this viewpoint as we believe the data show that international law and international norms can influence state behavior in a variety of circumstances.

have a preference for compliance with international law.¹⁴⁷ One weakness of this theory is that it lacks micro-foundations, which makes assessing how it works difficult.¹⁴⁸ That said, if an international instrument has a sort of gravitational pull, it is plausible that it could operate in both directions: just as states that otherwise perform below the international standard can be drawn upward, states that otherwise perform above this standard can be pulled down.

Drawing on Abram and Antonia Chayes and on Thomas Franck, Professor Harold Koh argues that treaties and non-binding international instruments can, under the right conditions, change the views of national government officials.¹⁴⁹ These officials interact with the international system, interpret it, and, ultimately, internalize its rules.¹⁵⁰ In political science, related viewpoints are most closely associated with the constructivist school of thought. Prominent constructivist theorists argue that following the adoption of an international human rights instrument, government leaders may be persuaded that their prior views were wrong, and their preferences may change as a result.¹⁵¹ Alternatively, leaders may be socialized to view certain actions as unacceptable to the international community and face peer pressure to modify their behavior.¹⁵²

If leaders are persuaded or socialized to expand protection for a particular right when international norms call for extensive protections, we believe state leaders can be similarly persuaded or socialized to limit their protections when the international norm calls for less stringent protections. Leaders might be persuaded, for example, that a lower level of protection for criminal defendants does not significantly increase the risk of placing innocent persons in jail while significantly reducing the costs of the administration of justice. Or even if leaders are not persuaded about the merits of this lower level of protection, they might simply follow their peers toward this new standard because they now consider it the appropriate level of protection for a modern state. Indeed, in their forthcoming book, Professors Ryan Goodman and Derek Jinks recognize this risk. They write, “Emulation and mimicry associated with acculturation can also produce a ‘race to the middle.’ The concern here is that states that

147. See Abram Chayes & Antonia Handler Chayes, *On Compliance*, 47 INT’L ORG. 175, 175–76, 178 (1993); THOMAS M. FRANCK, *THE POWER OF LEGITIMACY AMONG NATIONS* 195 (1990).

148. See LINOS, *supra* note 1, at 13–17.

149. See Harold Hongju Koh, *Why Do Nations Obey International Law?*, 106 YALE L.J. 2599, 2602, 2646 (1997) (reviewing ABRAM CHAYES & ANTONIA HANDLER CHAYES, *THE NEW SOVEREIGNTY: COMPLIANCE WITH INTERNATIONAL REGULATORY AGREEMENTS* (1995) and THOMAS M. FRANCK, *FAIRNESS IN INTERNATIONAL LAW AND INSTITUTIONS* (1995)).

150. *Id.*

151. See Jeffrey T. Checkel, *Why Comply? Social Learning and European Identity Change*, 55 INT’L ORG. 553, 577 (2003).

152. See Ryan Goodman & Derek Jinks, *International Law and State Socialization: Conceptual, Empirical, and Normative Challenges*, 54 DUKE L.J. 983, 984–85 (2005).

would otherwise aspire to heightened levels of human rights protection gravitate to lower expectations or standards of success.”¹⁵³

In short, a variety of theoretical mechanisms could lead to backsliding. We propose a particular theoretical model to clarify the logic of the argument, but alternative accounts of why states respond to international norms may lead to similar results.

V. IMPLICATIONS

Our theory of backsliding has important implications for states and others involved in designing human rights agreements. In this Section, we first apply the theory to highlight some practical implications for treaty membership and then describe how backsliding explains some prominent yet puzzling features of international human rights agreements.

A. Treaty Membership

We have discussed some of the implications of backsliding for human rights agreements, including the question of how strong the protections in agreements should be,¹⁵⁴ and the legal form they should take.¹⁵⁵ In this Section, we explore an additional set of implications concerning membership in human rights treaties.

Our theory provides one reason why the presence of high-performing states in a human rights agreement may result in more demanding human rights standards: higher standards reduce the likelihood of backsliding. Granted, these high-performing states presumably have a preference for high human rights standards, and this should have some upward pull on the content of an agreement. That said, it is hard to explain why a high-performing state would choose to expend political capital to increase the standard of a human rights agreement with which it is already in compliance. Absent an effect like backsliding, the agreement should have no direct effect on a state that is already in compliance with its terms.

One possible explanation is that a country prefers a higher standard because it cares about the welfare of citizens of foreign countries. Even when treaties are effective, however, it is not certain that higher standards lead to better human rights outcomes. In our model, and in many models of the relationship between international norms and domestic policy, only countries close to the international norm are pulled upward. As a result, as we explain

153. See RYAN GOODMAN & DEREK JINKS, *SOCIALIZING STATES: PROMOTING HUMAN RIGHTS THROUGH INTERNATIONAL LAW* 75 (2013).

154. See *supra* Part IV.B.1.

155. See *supra* Part IV.B.2.

above, it may be better to have a less high standard so as to pull up the worst-behaving states.¹⁵⁶

Backsliding provides an alternative (or perhaps additional) explanation. High-performing states do not seek higher standards out of a concern for the treatment of foreign individuals, or at least not only for this reason. They prefer a higher standard so as to reduce and perhaps eliminate backsliding. Even if a state is unable to raise the international norm to the high level it provides to its own citizens, the closer the norm gets to that level, the less backsliding the state is likely to face.

One implication of the above discussion is that having high performers at the table may distort the resulting agreement and lead to higher-than-optimal standards. More generally, the identity of the states involved will influence the content of the agreement. This is, of course, both obvious and well understood, but an appreciation of backsliding changes the way we think about the preferred mix of participants.

Once we realize that the identity of the states involved in norm-formation matters, it is natural to think about how one might manage participation to achieve better results. One point has already been mentioned—the appeal of regionalism. If states in a particular region are similar to one another in their human rights conduct, it is easier to create a norm that will have positive effects while minimizing backsliding. Put another way, a group of similar states committed to an improvement in their human rights regimes may generate a more valuable agreement if higher-performing states are not present or involved in the process.¹⁵⁷ This is because high-performing states may push for norms that are too high, in the sense that they do not affect enough low-performing states.

For human rights advocates, however, the absence of high-performing states can present its own disadvantages. For an agreement to generate an upward pull on existing policies requires, of course, that the resulting norm be more protective of human rights than the practices of the target states. A group of states that is more uniform in its behavior may not be enthusiastic about changing that behavior. Without a relatively high-performing state present, one wonders where the pressure for a high standard will come from.

We thus have the curious, and perhaps somewhat uncomfortable, result that an agreement capable of exerting an upward pull on lower-performing

156. See *supra* Part III.A.

157. There are many influences here, and we do not attempt to capture them all. The incentive to exclude high-performing states described above, for example, may be partially or fully offset by the fact that the presence of a more diverse set of states might give the norm a strong claim to universality and, therefore, make it more forceful. Because there are many influences that we cannot engage in this Article, the points made in the text should be viewed as additional factors, not the entire story. For a broader discussion of the advantages and disadvantages of variations in treaty membership, see generally Gabriella Blum, *Bilateralism, Multilateralism, and the Architecture of International Law*, 49 HARV. INT'L L. J. 323 (2008).

states may require the presence of at least some high-performing states. But having those high-performers at the table risks pulling the norm above the level that would yield the greatest improvements to human welfare. One possible strategy, then, is to have high-performing states encourage, or pressure, low-performing states to enter into agreements that ultimately they themselves do not join.¹⁵⁸ There are obvious reasons to be suspicious of such an approach, including the possibility that the high-performing states will impose requirements that serve their own interests rather than those of the participating states. For this reason, we do not advance this as a normatively desirable proposal. Rather, we raise it as one example of how an appreciation of backsliding can change our normative conclusions.

The issue of membership is still more complicated once we acknowledge that international norms do not have to be consent based. Norms may emerge from declarations of international organizations that do not require unanimous support (e.g., United Nations resolutions) or that do not include all states (e.g., OECD agreements). They may also be the result of agreements that are non-binding, such as the UDHR. The list of potential sources of an international norm is very long and includes treaties, non-binding agreements, customary international law, other states' practice, decisions of tribunals, and more. The possibility that international norms will develop without every state's consent is a feature of international human rights regimes generally, but it plays out differently with backsliding. If one ignores backsliding, non-consensual human rights norms (for better or worse) have the potential to exert an upward pull on low-performing states, even if those states are not involved in the generation of the norm or have not consented to it.

However, when we consider backsliding, the non-consensual aspect of human rights norms works in the other direction. A high-performing state that objects to a norm because it risks causing backsliding cannot avoid it simply by refusing to join a relevant treaty or organization; that state may be influenced by the norm even if it does not consent. If we continue to assume a correlation between powerful states and high-performing states, there is a risk that high-performing states will find ways to undermine norm-creation efforts in an attempt to avoid backsliding.

B. Some Empirical Puzzles Explained

Our theory of backsliding helps to explain three particularly puzzling patterns in international human rights regimes: the reluctance of top performers to join international agreements, the proliferation of regional human rights

158. We have previously discussed the fact that states may be affected by international norms even if those norms are not legally binding. That same possibility applies here, meaning that high-performing states cannot fully escape the risk of backsliding. By refusing to join the agreement, however, they weaken the transmission of the norm to their domestic systems and, therefore, reduce the associated risk. *See supra* Part IV.B.1.

systems, and the extremely high standards that characterize many human rights agreements.

1. The Behavior of High-Performing States

We start with the observation, made in several empirical studies of human rights agreements, that top human rights performers are often reluctant to join international agreements.¹⁵⁹ Professors Emily Hafner-Burton, Kiyoteru Tsutsui, and John Meyer note that while “it seems obvious that states with negative human rights records might be somewhat more reluctant than others to ratify treaties that subject them to intensified internal and external criticism[,] [t]he empirical data contradict this intuitive expectation.”¹⁶⁰ Similarly, Professor Oona Hathaway reports that “the states with the best human rights practices (and hence the best reputations) are often more reluctant to join human rights treaties than those with worse practices (and hence worse reputations).”¹⁶¹ This pattern is puzzling because under existing theories of how human rights instruments interact with domestic policy making, the cost of joining international agreements providing protections that a country already complies with should be very low.¹⁶²

Our theory of backsliding, however, is consistent with this pattern and predicts it.¹⁶³ That said, it is not the only explanation: several other factors could also account for the reluctance of high-performing countries to join international agreements.¹⁶⁴ Still, top performers might worry that the relevant international standards would exert a downward pull on their domestic standard and that joining a treaty would impede future progress or even cause a reversal of human rights gains. Indeed, sometimes advocates for minority groups oppose the ratification of international treaties intended for their benefit by making closely related arguments about backsliding. For example, advocates for Canada’s First Nations successfully lobbied against ratification of ILO

159. See Andrew Moravcsik, *The Origins of Human Rights Regimes: Democratic Delegation in Postwar Europe*, 54 INT’L ORG. 217, 228–29 (2000); see also SIMMONS, *supra* note 1, at 67–77 (arguing that governments that already respect rights can still face substantial domestic political costs when considering whether to join an international treaty).

160. Emilie M. Hafner-Burton, Kiyoteru Tsutsui & John W. Meyer, *International Human Rights Law and the Politics of Legitimation: Repressive States and Human Rights Treaties*, 23 INT’L SOC. 115, 117 (2008).

161. Oona A. Hathaway, *The New Empiricism in Human Rights: Insights and Implications*, 98 AM. SOC’Y INT’L L. PROC. 206, 209 (2004).

162. See George W. Downs, David M. Rocke & Peter N. Barsoom, *Is the Good News About Compliance Good News About Cooperation?*, 50 INT’L ORG. 379, 393–99 (1996).

163. See discussion *supra* Part V.A.1.

164. Oona A. Hathaway et al., *The Treaty Power: Its History, Scope, and Limits*, 98 CORNELL L. REV. 239, 305 (2013); see, e.g., DAVID SCHEFFER, *ALL THE MISSING SOULS* 166–67 (2012) (describing the United States’ reluctance to ratify the Rome Statute). Another reason for hesitancy from already-complying countries may be that many of the worst rights violators do ratify. See Hathaway, *supra* note 161, at 209; Oona A. Hathaway, *Why Do Countries Commit to Human Rights Treaties?*, 51 J. CONFLICT RESOL. 588, 588 (2007).

Convention 169 on Indigenous People. They worried that Canada would settle for the low standards of rights protected by the Convention, and that the Canadian government would later oppose increasing these standards by referencing its compliance with this Convention.¹⁶⁵

2. Regionalism

Our theory also helps explain a second puzzle in international human rights law concerning the strength and number of regional human rights systems. Though human rights practices vary from one region to another, supporters often emphasize the universal nature of human rights. Indeed, human rights claims derive much of their moral power from the belief that they are universal: that all humans deserve to enjoy certain fundamental rights.¹⁶⁶ The assertion of universality is also an essential part of the argument that demanding compliance with human rights norms is not an unjustified interference in the domestic affairs of the state.¹⁶⁷

The rhetoric of universality is somewhat incongruous with the fact that many of the most important human rights institutions are regional rather than universal. The ECtHR, for instance, has been called “the crown jewel of the world’s most advanced international system for protecting civil and political liberties,” reviewing tens of thousands of cases annually and protecting the

165. See, for example, the statements of Judith Sayers, a Nuu-Chah-Nulth lawyer from British Columbia, in her article titled “ILO Convention Must Be Stopped”:

If Indians in Canada ask the Canadian government to ratify this convention, it will be seen as a consent to a very low standard of rights. The Canadian government will use this as a knife in our backs as we continue our work internationally. They will maintain they do not have to raise the standard higher, because indigenous people agreed to the convention. That was the strategy behind revising the ILO convention. Now, many states are referring to it, stating this is what a Declaration of Indigenous Rights should look like.

Judith F. Sayers, *ILO Convention Must be Stopped*, 8 WINDSPEAKER, no. 2, 1990, at 4, available at <http://205.186.158.152/publications/windspeaker/ilo-convention-must-be-stopped-0>. The convention was criticized by several other indigenous groups as well, which successfully passed a resolution opposing its ratification. See, e.g., Sharon Venne, *The New Language of Assimilation: A Brief Analysis of ILO Convention 169*, 2 WITHOUT PREJUDICE: EAFORD INT’L REV. RACIAL DISCRIMINATION 53, 66–67 (1989); see also DOUGLAS SANDERS, DEVELOPING A MODERN INTERNATIONAL LAW ON INDIGENOUS PEOPLES 9, 14 (1994), available at <http://www.ubcic.bc.ca/files/PDF/Developing.pdf>. Related concerns about assimilation pressures had been raised earlier with respect to ILO Convention 107, the predecessor of ILO Convention 169, and may “have prevented its ratification by some more progressive countries which might otherwise have wished to apply its protective provisions.” Lee Swepston, *A New Step in the International Law on Indigenous and Tribal Peoples: ILO Convention No. 169 of 1989*, 15 OKLA. CITY U. L. REV. 677, 683 (1990).

166. See Jack Donnelly, *The Relative Universality of Human Rights*, 29 HUM. RTS. Q. 281, 282 (2007); see also JACK DONNELLY, UNIVERSAL HUMAN RIGHTS IN THEORY AND PRACTICE 7 (2d ed. 2003).

167. See generally Jack Donnelly, *State Sovereignty and International Intervention: The Case of Human Rights*, in BEYOND WESTPHALIA? STATE SOVEREIGNTY AND INTERNATIONAL INTERVENTION 115 (Gene M. Lyons & Michael Mastanduno eds., 1995).

rights of eight hundred million people.¹⁶⁸ Similarly, the Inter-American Court of Human Rights has great influence on Latin American governments' choices.¹⁶⁹ And the more recently established African Court of Human and People's Rights has a particularly wide-ranging mandate.¹⁷⁰ The importance of regional human rights systems in Europe, the Americas, and even Africa is puzzling as it is hard to develop forceful normative claims that derive from residence on a particular continent. It is easy to support the moral claim that every human being is endowed with certain rights, but much harder to support a similar moral claim that is limited to residents of Europe. Moreover, as prominent international lawyers have noted, regionalism has led to the proliferation of human rights tribunals, giving rise to important risks, such as conflict and contradictions in evolving human rights doctrine, and overlapping jurisdiction for individual cases.¹⁷¹

Our theory of backsliding helps explain this puzzle by suggesting that different standards for countries with different capabilities reduce the risk that top-performing countries will be affected by a low universal standard.¹⁷² If one assumes that human rights are universal and that creation or clarification of international norms serves only to improve the behavior of low-performing states, a more sensible policy would be to establish a single set of global norms. If one were concerned that establishment of highly protective norms would fail to influence low-performing states, the solution is not regionalism but tiered human rights norms. That is, one could identify the "true" human rights norm (by assumption a highly protective one) along with a series of lower "intermediate" norms. States that managed to meet the intermediate thresholds would be recognized as falling short of the ideal, but making progress.

The theory of backsliding helps explain why states have opted for regional human rights norms rather than limiting themselves to universal norms. To the extent backsliding is a concern, regional norms can be tailored to achieve maximum benefit for the affected countries. The maximization of benefits must account for both the potential to generate an upward pull on low performers and the risk of causing backsliding by high performers. A regional norm is crafted for a set of countries that feature less diversity in human rights conduct than would be the case for a global norm. If, for example, countries in the region are high performers from a global perspective, then regional norms can

168. Laurence R. Helfer, *Redesigning the European Court of Human Rights: Embeddedness as a Deep Structural Principle of the European Human Rights Regime*, 19 EUR. J. INT'L L. 125, 125–26 (2008).

169. See Lea Shaver, *The Inter-American Human Rights System: An Effective Institution for Regional Rights Protection?*, 9 WASH. U. GLOBAL STUD. L. REV. 639, 647–49 (2010).

170. See Cesare P.R. Romano, *The Proliferation of International Judicial Bodies: The Pieces of the Puzzle*, 31 N.Y.U. J. INT'L L. & POL. 709, 722 (1999).

171. See Benedict Kingsbury, *Foreword: Is the Proliferation of International Courts and Tribunals a Systemic Problem?*, 31 N.Y.U. J. INT'L L. & POL. 679, 681–84 (1999).

172. See discussion *supra* Part V.A.2.

be ambitious. Norms set at a high level will not greatly exceed the current levels set by individual states, so there is reason to expect that the norms will exercise some upward pull. Demanding norms will simultaneously reduce the risk of backsliding because they will still be more protective than the existing standards of many states in the region.

A global norm, on the other hand, would impose higher backsliding costs. This description is consistent with what we observe within the European human rights system as well as in the human rights system in the Americas where “the Inter-American Court of Human Rights . . . [has] perhaps imposed higher standards on states than has the ICJ” on questions of state responsibility to take affirmative measures of protection.¹⁷³

If a region is populated by relatively low-performing states, less ambitious norms will be more effective at generating upward pull and will have only modest backsliding costs.¹⁷⁴ A global standard that is set at a high level, on the other hand, will likely be ignored by war-torn and impoverished states whose citizens are in need of just such protections.

3. Aspirational Norms

A third puzzle that our theory explains concerns the very high standards that human rights treaties often enshrine. In the non-human rights context, other international agreements, including many military, trade, and environmental agreements, set modest standards that impose only small demands on participating states.¹⁷⁵ In the human rights arena, in contrast, international agreements frequently set very high standards, “above a level that many participating countries can or want to comply with immediately or within the foreseeable future.”¹⁷⁶ For example, according to critics, several “CEDAW provisions are plainly unrealistic . . . [w]ith the gross burden of providing basic needs to their citizens still unfulfilled, many developing nations hesitate to take on the additional responsibility of un-stereotyping women’s and men’s roles, as Article 5 of the convention suggests.”¹⁷⁷ International treaties that call for social and economic rights are similarly criticized for setting impossibly high standards. For example, Article 11 of the ICESCR calls for “the right of

173. See Kingsbury, *supra* note 171, at 861.

174. Indeed, the follow-on protocols of the ECtHR, the Inter-American Court of Human Rights, and the African Court of Human Rights contain more stringent rights protections than their original treaty formations and equivalent global conventions. This model of a primary convention with supplementary protocols may function to constrain backsliding. See discussion *supra* Part IV.B.3. It is less clear whether overlapping international laws on the same subject matter (e.g., violence against women, children’s rights, disappearances, discrimination, and protection of people with disabilities) impede or aid backsliding.

175. See Downs, Roche & Barsoom, *supra* note 162, at 388–92.

176. Neumayer, *supra* note 8, at 928.

177. Meghana Shah, *Rights Under Fire: The Inadequacy of International Human Rights Instruments in Combating Dowry Murder in India*, 19 CONN. J. INT’L L. 209, 222 (2003).

everyone to an adequate standard of living for himself and his family, including adequate food, clothing and housing,” a right that many developing countries cannot easily fulfill.¹⁷⁸

Critics worry that very high standards dilute the status of human rights agreements as legally binding obligations and limit their potential to influence states. For example, philosopher Onora O’Neill worries that such aspirational standards dilute the value of human rights generally, and lead us to “accept that where human rights are unmet there is no breach of obligation, nobody at fault, nobody who can be held to account, nobody to blame and nobody who owes redress.”¹⁷⁹ It is not unusual for states to enter into human rights agreements that all involved expect will be routinely violated. Nowhere is the disconnect between promise and behavior as stark as in the case of human rights.

Our backsliding theory helps explain why this curious acceptance of empty promises exists in human rights agreements more so than elsewhere. If states are concerned about backsliding, they have an incentive (all else equal) to prefer high standards over low standards in a way that is not explained by existing theories of international human rights. A demanding norm—perhaps even one that exceeds every country’s existing practice—reduces the associated backsliding costs because it reduces the number of states for which the international norm is less demanding than their *ex ante* domestic policy. In the most extreme cases, such as the CEDAW requirement that states take “all appropriate measures . . . to modify the social and cultural patterns of conduct of men and women,” the number of such countries may be reduced to zero, which eliminates the risk of backsliding completely.¹⁸⁰ While very high standards reduce the upward pull of human rights agreements on low performers, they also reduce the risk of backsliding among top performers.

In this Section, we have explored important implications of our theory of backsliding for the design of human rights regimes. In particular, our theory helps explain some puzzling features of these regimes, including the reluctance of top performers to join international agreements whose standards they already meet, the proliferation of regional human rights systems, and the very high standards that human rights agreements often specify. Our theory also opens up new questions about the ideal design of human rights treaties, including questions about participation and treaty membership.

178. International Covenant on Economic, Social and Cultural Rights, *supra* note 26, at art. 11.

179. See Onora O’Neill, *The Dark Side of Human Rights*, 81 INT’L AFF. 427, 430 (2005).

180. Convention on the Elimination of All Forms of Discrimination Against Women, *supra* note 25, at art. 5(a).

CONCLUSION

Human rights violations remain a widespread phenomenon. International law in general, and international human rights law in particular, seek to influence the ways that states treat their citizens.

The study of human rights law is, in part, the study of how international legal rules can be used more effectively to advance this objective. As a normative matter, backsliding may be seen as an undesirable feature of the international system, but that in no way diminishes the importance of understanding how it works. The stakes involved are too great to allow complacency with respect to our study. We must seek the best possible understanding of how international norms affect behavior, even if at times they do so in undesirable ways.

This Article makes it clear that we should not simply assume that when human rights standards improve state practices in some parts of the world, they always allow other countries—countries that currently perform above those standards—to continue on their upward trajectories. As we have shown, there is a risk that low international standards will empower domestic opponents of human rights and lead to a reduction of rights in high-performing countries.

It is not our purpose to undermine any ongoing human rights efforts, and we do not believe this Article does so. We believe, rather, that an understanding of backsliding allows both advocates and policy makers to think more clearly about how human rights norms affect the outcome that really matters: the way in which humans are treated.

Backsliding is a novel argument in the human rights literature. In this Article we have aimed to establish that backsliding is theoretically plausible, explore the conditions that affect its magnitude, and identify some tools that might mitigate its consequences. However, much work remains to be done. Some of this work will hopefully be empirical, and will help identify the size of the problem as well as the issue areas and countries where backsliding poses the greatest risk. Other work, we hope, will be pragmatic and policy oriented, concerned with identifying the best ways to design international standards, draft international instruments, and organize advocacy campaigns with an eye toward improving human rights in some countries while mitigating the risk of backsliding in others.