2015

National Human Rights Institutions: Adoption, Design, and Effectiveness

Ryan M. (Ryan Merrill) Welch
NATIONAL HUMAN RIGHTS INSTITUTIONS:
ADOPTION, DESIGN, AND EFFECTIVENESS

By
RYAN M. WELCH

A Dissertation submitted to the Department of Political Science in partial fulfillment of the requirements for the degree of Doctor of Philosophy

2015

Copyright © 2015 Ryan M. Welch. All Rights Reserved.
Ryan M. Welch defended this dissertation on July 9, 2015.
The members of the supervisory committee were:

Will H. Moore  
Professor Co-Directing Dissertation

Christopher Reenock  
Professor Co-Directing Dissertation

Sumner Twiss  
University Representative

Quintin H. Beazer  
Committee Member

Megan Shannon  
Committee Member

The Graduate School has verified and approved the above-named committee members, and certifies that the dissertation has been approved in accordance with university requirements.
I wish to offer my sincerest gratitude to Will Moore. He offered the perfect proportion of challenge and empathy that allowed me to stay focused on the dissertation. All of my graduate successes – not just this dissertation – stem from Will’s uncanny skill and ceaseless dedication as an advisor. Thanks also to my committee, who offered insightful feedback on my work throughout the process. Special thanks to Quintin Beazer, who went above and beyond as a committee member. He often popped into my office just to see how I was handling the stresses of graduate school. His obvious care led me to seek him for advice throughout the writing process. Special thanks to my colleagues and friends who suffered, laughed, and ultimately learned with me throughout graduate school – especially Jack Mewhirter, Eryn Jones, and Scott Meachum. Thanks to Rachel for providing emotional support as I dealt with the whirlwind that is finishing the dissertation. Lastly, I’d like to thank my family for always believing in my abilities as an academic – especially Mom and Chet, both of whom were integral to my ability to finish. Chapter 1 is forthcoming in *Journal of Human Rights*. 
# TABLE OF CONTENTS

List of Tables ................................................................. vi
List of Figures ................................................................. vii
Abstract ...................................................................... viii

1 National Human Rights Institutions: Domestic Implementation of International Human Rights Law 1
   1.1 International Law and Torture ......................................... 3
   1.2 National Human Rights Institutions ................................... 5
   1.3 The Importance of NHRIs and Information on Compliance ........... 6
      1.3.1 Legal Redress ..................................................... 7
      1.3.2 Mobilization ......................................................... 9
   1.4 Data and Methods .......................................................... 11
      1.4.1 Dependent Variable ............................................... 11
      1.4.2 Explanatory Variable .............................................. 12
      1.4.3 Control Variables ................................................ 12
      1.4.4 Statistical Model ................................................. 15
   1.5 Results ................................................................. 15
   1.6 Conclusion ............................................................ 19

2 The Role of Reputation in the Norms Cascade: The Case of NHRI Adoption 22
   2.1 The Liberal Order and NHRIs .......................................... 24
   2.2 NHRI Adoption .......................................................... 25
   2.3 Past Explanations ....................................................... 26
   2.4 Diffusion, Reputation, and Status ..................................... 27
   2.5 Importance of Shaming ............................................... 28
   2.6 Research Design ........................................................ 30
      2.6.1 Spatial-Temporal Domain ....................................... 30
      2.6.2 Model ............................................................... 30
      2.6.3 Variables .......................................................... 30
   2.7 Results ................................................................. 38
   2.8 Conclusion ............................................................. 40

3 Delegating Accountability: An Analysis of National Human Rights Institutions Variation 43
   3.1 A Brief History of NHRIs .............................................. 45
   3.2 Explaining Variation in NHRI Accountability Mechanisms ............. 46
   3.3 Research Design ........................................................ 52
      3.3.1 First Stage Variables ............................................. 53
      3.3.2 Second Stage Variables .......................................... 55
      3.3.3 Control Variables ................................................ 57
   3.4 Results ................................................................. 59
Appendix

A Supplementary Material for Chapter 2 65
A.1 Latent Variable Models 65
A.2 State Dominance Figures 66

B Supplementary Material for Chapter 3 68
B.1 Convergence Statistics 68
B.1.1 Trace Plots 68
B.1.2 Autocorrelation Plots 68
B.1.3 Geweke–Brooks Plots 68
B.2 Other Models 77

References 78
Biographical Sketch 95
LIST OF TABLES

1.1 Effectiveness Ordered Logit Results, 1981–2007 ........................................ 16
1.2 Probability of Changing Torture Categories for CAT Ratifier when It Adopts an NHRI 17
2.1 Adoption Control Variables ................................................................. 34
2.2 Effect of Shaming on NHRI Adoption, 1991–2009 ................................. 38
3.1 Design Independent Variables ............................................................ 57
3.2 Design Control Variables ................................................................. 59
3.3 Delegation of NHRI Accountability Mechanism, 1991-2012 ...................... 60
3.4 Design First Differences ................................................................. 62
B.1 Design Model with No Uncertainty ..................................................... 77
B.2 Design Zelig Results ................................................................. 77
# LIST OF FIGURES

1.1 States that Ratified the CAT by 2007. Source: Author’s calculation from sources described in *Explanatory Variables* section. .............................................. 13

1.2 Effect of Interaction Term and Constituents (with 95% Confidence Intervals) .... 16

2.1 Global Adoption of NHRI’s by year ............................................. 31

2.2 Fifty Most Shamed States ............................................................ 32

2.3 Fifty Least Shamed States ............................................................ 33

2.4 Hazard Function of Adopting an NHRI, 1991-2009 ............................. 37

3.1 How NHRI’s are Established (based on data from Conrad, DeMeritt and Moore, 2013) 51

3.2 NHRIs that can Levy Punishment, 2012. White=No NHRI, Gray=NHRI, no Punishment, Black=NHRI, Punishment .................................................. 54

3.3 Permitted Activities of NHRIs. Source: (Conrad, DeMeritt and Moore, 2013) .... 56

A.1 Fifty Most Materially Dominant States .......................................... 66

A.2 Fifty Least Materially Dominant States .......................................... 67

B.1 Trace Plots .................................................................................... 69

B.2 Trace Plots II ............................................................................... 70

B.3 Autocorrelation Plots .................................................................... 71

B.4 Autocorrelation Plots II ................................................................. 72

B.5 Autocorrelation Plots III ............................................................... 73

B.6 Geweke–Brooks Plots ................................................................. 74

B.7 Geweke–Brooks Plots II ............................................................... 75

B.8 Geweke–Brooks Plots III ............................................................... 76
ABSTRACT

Does adopting a National Human Rights Institution (NHRI) make states’ international commitments more constraining? If so, why would states adopt the constraining institution in the first place? When states do adopt an NHRI, how do they design it? I answer these three questions in three separate essays in the dissertation. Past scholarship emphasizes the role of domestic institutions for international human rights commitments. NHRI s, domestic institutions tasked with the protection and promotion of human rights, represent another institution to explore. Focusing on the generally secretive practice of torture, I argue NHRI s provide information to potential mobilizers and domestic legal systems increasing the probability states follow through on their commitment not to torture. I find that when a country already ratified the Convention Against Torture (CAT), the presence of an NHRI substantively decreases the chances the state tortures often. Given NHRI ability to constrain state behavior, why do states adopt NHRI s? Most NHRI adoption happens within a norms cascade. I argue states weigh the reputational benefits with the costs of NHRI adoption. Shaming by other actors approximately signals how those actors perceive the state in question’s status with respect to the international society of states. Those states never or always shamed do not reap high enough benefits to adopt; more adoption exists in the middle. I test the hypothesis with a Cox Proportional Hazards model and find support that states adopt NHRI s to manage their international status. Not all NHRI s are created equally, though. Considerable variation exists pertaining to what actions NHRI s can legally take to affect their charge; some even levy punishment. The final essay proposes a theory of legislature delegation to NHRI s with the executive seeking to influence the process to retain maximum discretion in the future. Electoral laws and political situations allow the executive to leverage her power with respect to the legislature in order to decrease the probability the NHRI may levy punishment. The executive may even hijack the whole process and create the NHRI unilaterally through decree also decreasing the probability of legal NHRI punishment powers. Estimating a Bayesian logistic regression, that accounts for selection and uncertainty, I find support for the theory.
CHAPTER 1

NATIONAL HUMAN RIGHTS INSTITUTIONS: DOMESTIC IMPLEMENTATION OF INTERNATIONAL HUMAN RIGHTS LAW

Does the presence of an NHRI make the UN Convention Against Torture (CAT) effective at reducing state torture\(^1\) practices? Human rights have been codified into international law since the founding of the United Nations (1945), and especially after the Universal Declaration of Human Rights (UDHR) (1948). Treaties represent the main instrument of codification. A number of scholars explore these treaties in order to judge just how effective they achieve, as the UDHR Preamble calls for, “the promotion of universal respect for and observance of human rights and fundamental freedoms.” Scholars have found mixed results for how effective international treaties are at inducing compliance. Studies of multiple treaties find correlations between ratification and compliance for some of the treaties (e.g Hathaway, 2002; Hill, 2010; Simmons, 2009). However, much of the news remains negative, as many scholars find those states that ratify human rights treaties do not change their behavior (Hafner-Burton and Tsutsui, 2005; Hathaway, 2002; Simmons, 2009), or even perpetrate worse behavior (Hathaway, 2002; Hill, 2010; Hollyer and Rosendorff, 2011; Simmons, 2009; Vreeland, 2008). Those studying international law’s effect on state torture started by evaluating signature/ratification (e.g Hathaway, 2002) only to quickly recognize the importance of domestic politics and institutions (e.g Hollyer and Rosendorff, 2011; Lupu, 2013; Powell and Staton, 2009; Simmons, 2009; Vreeland, 2008).

I contribute to the study of domestic factors affecting international institutional effectiveness by proposing another domestic institution with the ability to make international human rights commitments effective – national human rights institutions (NHRIs). As such, this study rep-

\(^1\) I define torture as does the CAT: “any act by which severe pain or suffering, whether physical or mental, is intentionally inflicted on a person for such purposes as obtaining from him or a third person information or a confession, punishing him for an act he or a third person has committed or is suspected of having committed, or intimidating or coercing him or a third person, or for any reason based on discrimination of any kind, when such pain or suffering is inflicted by or at the instigation of or with the consent or acquiescence of a public official or other person acting in an official capacity” (http://www.un.org/documents/ga/res/39/a39r046.htm)
resents the first quantitative study, known to this author, proposing NHRI effectiveness with respect to international treaties with the specific mechanism of information raising.\(^2\) NHRIs are domestic institutions established by the government to be watchdogs on the government with respect to human rights practices. They provide information to domestic and international audiences about legal obligations and compliance that ultimately affects actors’ pursuit of legal redress and mobilization. Both legalization and mobilization bring potential costs to the executive in the form of decreased rents and/or increased probability of losing power. In expectation of these costs, the executive should torture less often. Consider the recent “Wheel of Torture” case in the Philippines (Balana, 2014; Calamur, 2014). The NHRI in the Philippines, the Commission on Human Rights (CHR), investigated a supposed secret prison and found a hidden detention cell that contained a multi-colored wheel with different torture tactics that was spun in a game of “torture roulette” by officers to be enacted on prisoners (Amnesty International, 2014). The CHR identified 41-44 victims and worked to disseminate their findings by sharing information with various news outlets (including the Philippine Daily Inquirer and the Associated Press) and INGOs such as Amnesty International (Amnesty International, 2014; Calamur, 2014). Various rights groups are calling for [President] Aquinos’ position due to “violations [that] have continued with impunity under [his] watch” despite legal obligations, including acceding to the UN CAT in 1986 (Balana, 2014). The CHR supplied information to various interested actors that resulted in the (threat of) mechanisms discussed later in the paper – legal redress and mobilization.

Most studies include a number of different human rights treaties covering different abuses from genocide to freedom of speech, but rely on theorizing about human rights in general which can be misleading due to the different underlying causes for different types of abuses. I contribute to the scholarship that highlights specific abuses by focusing exclusively on freedom from torture. Why focus on torture? Torture is a “dastardly and totally inhuman act” (Filartiga v. Pena-Irala, 630 F.2d 876, 883; quoted in Luban, 2006). It is the first explicit act condemned in the UDHR behind slavery. Despite the abhorrence towards torture, though, it is the most frequently violated physical integrity right in the last quarter of the twentieth century (Cingranelli and Richards, 1999, 522); so much so some refer to torture as a “normal” tool of statecraft (Conrad and Moore, 2010, 474).

\(^2\) The only other quantitative piece on NHRI effectiveness known to this author (Cole and Ramirez, 2013), describe NHRIs as a consequence of the global polity’s concern with certain rights (physical integrity) over time.
Torture is important due to reasons of normative concern and widespread usage. More importantly, the secretive way torture is perpetrated is especially suited for the mechanism proposed in the paper – increased information by NHRIs. Torture is a unique act governments perpetrate to compel testimony, punish (perceived) dissidents, and/or cow (sub)populations (e.g. Conroy, 2000; Rejali, 2007). While leaders may have common motivations for violating other rights, they are not the same set of motivations. Historically, judges ordered suspected criminals tortured, often in public (Peters, 1985). However, modern torture is used for confession, information, or intimidation, and given the norm against torture, is usually perpetrated in private (Rejali, 2007). The ability of NHRIs to shed light on the secretive practice is noteworthy.

The rest of the study is laid out as follows. In the next section I trace what scholars have contributed thus far about the effect of international law on state torture practices, eventually exploring the domestic institutions that make the law effective. I’ll then introduce NHRIs as important to international torture law compliance. These sections lead me to a testable hypothesis – states that have ratified the CAT torture less often when NHRIs are present. The following section lays out the research design, data, and methods used to test the hypothesis. I find support for the hypothesis estimating statistical models on 153 countries over the years 1981–2007: when states party to the UN CAT have an NHRI, they torture less often than if they did not have an NHRI. Finally, I conclude with implications for policy and future research.

1.1 International Law and Torture

Due, at least partly, to the memory of Nazi torture during World War II, torture was a priority of the nascent international human rights regime (Rejali, 2007); so much so, that Rodley (1999) considers rules against torture to have a privileged status in international law. The International Covenant on Civil and Political Rights (1966) lists freedom from torture among a comprehensive list of rights. In 1984, the UN General Assembly adopted the UN Convention against Torture and Other Cruel, Inhuman or Degrading Treatment or Punishment (CAT) specifically prohibiting torture. The CAT defines torture and how states should implement the convention (including domestic legislation, reporting, and monitoring). According to the UN Treaty Collection, to date, 154 states have either signed or ratified the CAT.

3 Although Rejali (2007) reminds us that despite having the reputation for innovative sadism, Nazi Germany borrowed many of its torture techniques from other cultures (e.g. electrotorture from the French).
Studies exploring the effect of international law on state torture practices yielded initially surprising if not depressing results (e.g. Hathaway, 2002; Hill, 2010). Both Hathaway (2002) and Hill (2010) find that ratification of international agreements aimed at curbing torture is associated with more state torture. Hathaway (2002) attributes the finding to the “expressive” role of committing to treaties, in which states select into the treaty treatment based on their international reputation needs. Hill (2010) uncovers the effect of treaties as exogenous treatments in order to deal with the endogeneity issue using a matching algorithm. He finds the CAT still associated with higher torture levels.

International human rights scholars started theorizing and testing conditional effects of international legal instruments (e.g. Lupu, 2013; Moravcsik, 2000; Powell and Staton, 2009; Simmons, 2009). Though the CAT isn’t associated with better torture practices on its own, the domestic political context offers opportunities for treaty success. For instance, using experiments, Wallace (2013) finds a state being party to international instruments forbidding torture lead to the state’s population being less supportive of torture. The finding is important because the population, upon attaining knowledge of state torture, may punish the violating government, either through elections or contentious activity. In fact, Simmons (2009) finds the spaces that allow for political participation where grievances occur (i.e. weak and transitioning democracies) are where international human rights treaties, including the CAT, are most effective. Even if states sign onto the CAT as a tactical concession without intention to comply (Hafner-Burton and Tsutsui, 2005; Hathaway, 2002; Risse and Sikkink, 1999), if the political space is open enough, citizens and NGOs will hold governments to account, thus raising the cost of violating the agreement (Risse and Sikkink, 1999; Simmons, 2009).

An effective judiciary also allows international agreements to reduce torture in a given state (Hathaway, 2007; Lupu, 2013; Powell and Staton, 2009; Simmons, 2009). Once a country commits to a treaty, the judiciary is often charged with interpreting the law (Hathaway, 2007). In fact, Article 4 of the CAT requires each party to make torture punishable under domestic law. For instance, Canada has amended its Criminal Code to make torture punishable by fourteen

---

4 Although see Fariss (2014) for the argument that these depressing results are due to the changing standard of accountability influencing the way in which we measure torture.

5 Including the CAT, American Torture Convention, and the African Charter

6 Some specifications in Hathaway (2002) show no statistical relationship between ratification and practice

7 Although see Conrad and Moore (2012) for an argument that elections may not lead to a decrease in torture because marginalized groups (not a priority of the majority) are often targeted.
years’ imprisonment (Moyo, 2000). This process is much smoother in states with monist legal traditions, in which international law is incorporated right into the domestic legal structure. However, judiciaries in states with dualist legal traditions take into account international law not yet incorporated into domestic law when ruling (Sloss, 2009; Killander, 2010). So regardless of legal tradition, an executive must consider how a judiciary will rule on her actions. If she believes she may face domestic legal ramifications, she may not order agents to torture, or may work to keep agents from torturing.

1.2 National Human Rights Institutions

National human rights institutions (NHRIs) represent another institution able to make international human rights treaties such as the CAT more effective on the ground. NHRIs are domestic institutions created by the government of the state in which they act charged with promoting and protecting human rights. They are a continuation and growth from the ombudsman model that has its roots in 19th century Sweden. Ombudsman offices, originally conceived, do not have an explicit human rights mandate. Instead, they raise awareness of government misconduct and maladministration by allowing citizens to register complaints (Cheng, 1968). This model of government accountability spread throughout Scandinavia (Finland, 1919; Denmark, 1955; Norway, 1962) to other parts of the world (New Zealand, 1962; Australia, 1977) (Koo and Ramirez, 2009).

As individual human rights protection became increasingly internationalized through the UN and its treaty system, and since human rights violations are often perpetrated by governments, these ombudsman offices became the first NHRIs. The UN promoted NHRIs, recognizing the potential for domestic institutions to facilitate international human rights commitments (Mertus, 2009). Not surprisingly, then, ombudsman offices began collecting more information on human rights violations as a sort of government malpractice, either themselves or in collaboration with other institutions (Hill, 1974). The UN’s call for domestic implementation of international human rights standards continued to influence states into the 1990s culminating in the Paris Principles.

---

8 Monists treat international and domestic law as one legal system with international law occupying a hierarchically superior position. Dualists treat domestic and international law as two separate spheres whereby international law is applied domestically to the extent that it is incorporated into the domestic law (von Glahn and Taulbee, 2013, 118). For a more comprehensive review of monism and dualism see Marian (2007).
The Paris Principles were defined at the first International Workshop on National Institutions for the Promotion and Protection of Human Rights held in Paris in October 1991. They called for the establishment of NHRIs and guidelines to make them effective.

Many states tasked an existing domestic institution to fulfill the role of NHRI. These included the ombudsman offices described above, institutions born of conflict transition (e.g. Northern Ireland, South Africa), or other human rights organizations (e.g. Switzerland) (Smith, 2006). New NHRIs were also established in the last few decades as states looked to democratize and/or become legitimate members of the international community (Reif, 2000). Currently there are over 100 NHRIs throughout the world (Koo and Ramirez, 2009).

Many of the newly established NHRIs took on a different organizational structure than the ombudsman offices. They consisted of a group of decision makers that formed a commission explicitly charged with promoting and protecting human rights, commonly referred to as human rights commissions. Other states established institutions similar to ombudsman in structure, but differing in their explicit human rights focus (often called human rights ombudsman offices). These different types of NHRIs perform a number of tasks that collect and disseminate information about state human rights practices such as fielding complaints, visiting sites such as prisons, advising and drafting legislation, working with NGOs, educating the public, and publishing findings (e.g Carver, 2004; Pegram, 2012; Reif, 2012). Although NHRIs represent a heterogeneous group of institutions, it is important to recognize that they may have an average effect on state torture practices.

1.3 The Importance of NHRIs and Information on Compliance

This section lays out the argument for the importance of NHRIs with respect to states complying with the CAT. Before expounding, though, I list the assumptions I employ throughout the theory. I assume the executive wishes to remain in power in order to continue to collect rents associated with office. Acting rationally, she will employ whatever means increase her chances to do so. The executive’s preference for office and willingness to employ necessary tactics to keep

---

9 The Paris Principles were adopted in by the United Nations General Assembly in 1993.
10 Studies, such as Cole and Ramirez (2013) suggest there is reason to study them as one homogenous group — “...all types of human rights institutions improve long-term physical integrity outcomes but not civil and political rights practices” (I, emphasis added). As noted in the conclusion, future study should explore further what aspects of NHRIs are most effective at curbing rights violations, and when.
that office matter because I also assume she is the actor that is ultimately responsible for state’s repressive actions. I assume individuals, as (possible) victims of torture, prefer to live life free of pain; that is, individuals would prefer to not be abused by the state.

International agreements, such as the CAT, are important for those seeking redress from abuse because they act as lynchpins with which to make arguments in the political arena (Powell and Staton, 2009; Risse and Sikkink, 1999; Simmons, 2009). But, in order for international agreements to be used in this way, information about violations of rights must be known; that is, people that know their rights exist and people that know their rights have been violated are more likely to have grievances about how they were treated and to act upon them. In this way, NHRIs are domestic institutions that provide information to domestic and international audiences about legal obligations (i.e. the CAT) and compliance that ultimately effects actors’ pursuit of legal redress and mobilization. Both the legalization and mobilization bring potential costs to the executive in the form of decreased rents and/or increased probability of losing power. In expectation of these costs, the executive should torture less often. In the sections that follow, I explain how NHRIs act as information clearing houses that may increase legal redress or mobilization (or both).

1.3.1 Legal Redress

Torture victims usually belong to marginalized groups such as criminals, dissidents, and minorities (Conrad, Haglund and Moore, 2013; Conrad and Moore, 2012; Conroy, 2000; Peters, 1985; Rejali, 2007). Persons belonging to these groups often lack the resources and knowledge to take legal action (Meeker and Dombrink, 1992; Willging, 1968). Many NHRIs run education and awareness campaigns to inform citizens of their rights under international law and what to do if a right has been breached, including lodging a complaint with the NHRI or bringing suit against the state. For instance, the National Human Rights Commission of the Republic of Korea launched a campaign in which 3,043 citizens visited; ultimately educating them on their international legal rights and how to pursue them (NHRCRK, 2010). NHRIs offer victims a place to air grievances apart from the state apparatus that perpetrated their torture. Once an NHRI fields a complaint, it may investigate and decide which action to take to settle the grievance. The NHRI sometimes decides to send the case to the court, informing the citizen of his/her rights and helping them through the legal process (Carver, 2012; Moyo, 2000). For instance, in Zambia in 1998, two complainants, one beaten by police to the point of permanent hearing
loss and another tortured as a suspect of aggravated robbery, were referred to the courts by the Permanent Human Rights Commission (Zambia’s NHRI) (Moyo, 2000). Scenarios such as these increase the probability the executive is sanctioned.

NHRIs can also increase the chances that these legal sanctions happen by supplying information to the legislature. Article 4 of the CAT states “each state party shall ensure that all acts of torture are offences under its criminal law.” Legislatures delegate complex legal topics to specialists (Lupia and McCubbins, 1994). NHRIs devote their resources to human rights promotion and protection making them ideal specialists to consult. Many NHRIs audit legislation before passage, or even write legislation themselves to ensure the law protects citizens’ rights by enshrining international legal obligations into the domestic system (Carver, 2004, 2012). Laws must exist in order to be leveraged for legal redress. NHRIs help create laws and ensure existing laws address human rights concerns. The Malaysian Human Rights Commission (SUHAKAM) went as far as to audit every law on the books to reconcile it with international human rights instruments (Carver, 2004). While Malaysia is not party to the CAT, other NHRIs may use their example to square domestic and international law.

Legal sanctions that are a result of increased information can include monetary or political costs. Both decrease the probability of holding onto office. Monetary costs represent a direct drain on the precious rents extracted at earlier time due to her position. An NHRI need not direct a case to the judiciary system in order to enact monetary costs on the executive, either. Sometimes, the NHRI will make recommendations to the government to compensate victims monetarily. For instance, in 2009, Uganda recommended monetary rewards to 9 victims of torture averaging around 20,000,000 Ugandan shillings11 (UHRC, 2009).

Other states, individually or collectively, can also seek legal redress. The literature on compliance with international agreements is too large to review here,12 but certain points are worth highlighting. Skeptics believe international agreements such as the CAT seldom require change in state behavior (e.g. Downs, Rocke and Barsoon, 1996; Mearsheimer, 1995). However, with increased interdependence between states has come repeated encounters where international commitments can be used by states to signal their intentions and build reputations over time (e.g. Axelrod, 1984; Guzman, 2008; Keohane and Nye, 1979). In order for commitment to amount

---

11 Approximately 8,000 USD
12 See Simmons (1998) for a helpful review that is but a starting point.
to more than just cheap talk, though, states must have information on how well other states are complying with these commitments. The CAT holds states responsible for reporting on their own behavior. By delegating authority to NHRIs, the information can be more credible about state behavior. Although states may reap some benefits for making their commitments more credible, the independent nature of NHRIs\textsuperscript{13} cause them to hand over information about state practices that the state would want to hide. Although states may try to hide their torture practices (Rejali, 2007), NHRIs are often tasked with fulfilling reporting requirements to treaty committees, such as the UN Committee against Torture. In fact, the Operational Protocol for the Convention against Torture (OPCAT) requires a National Preventative Mechanism (NPM) which is usually fulfilled by the state NHRI (Carver, 2010).\textsuperscript{14} The more reliable information states have about states’ torture actions, the less likely an executive will torture due to the anticipated costs imposed by other states, whether repuational (e.g Guzman, 2008) or monetary (e.g Hafner-Burton, 2005; Lebovic and Voeten, 2009).

\textbf{1.3.2 Mobilization}

International treaties, such as the CAT, can also increase the probability of domestic mobilization, but information is not passed directly from the international institution to the populace (Simmons, 2009, 143–144). Information provided by NHRIs serve as “conceptual frames that may serve to animate the demands of those whose ability, regularly and at low cost, to turn their leaders out of office is much less secure” (Simmons, 2009, 144). Domestic mobilization can directly lead to a higher probability that a leader is turned out of office.\textsuperscript{15} Domestic mobilization can also lead to a “boomerang effect” wherein domestic groups get their message out to the global human rights polity\textsuperscript{16} who in turn put pressure on the executive (Risse and Sikkink, 1999, 19). These pressures often come in the form of economic sanctions (e.g Blanton and Blanton, 2007; Hafner-Burton, 2005; Lebovic and Voeten, 2009) which are taken from earlier-collected rents and

\textsuperscript{13} Although the Paris Principles call for independent NHRIs, a number of case studies show independence is not a binary concept and some NHRIs are more or less independent than others (e.g. Mertus, 2009). Skepticism about NHRI independence (e.g. Rosenblum, 2012) actually provides for a conservative test of my hypothesis, as I am looking for the average effect and those NHRIs that are less independent would dampen the overall effect.

\textsuperscript{14} Note the tests in this study do not directly address OPCAT. The temporal range of the data go to 2007, and OPCAT did not go into force until 22 June 2006. Future studies should look further into OPCATs relationship with NHRIs.

\textsuperscript{15} This could be through a continuum of mechanisms from democratic elections to revolution. The most important point is domestic mobilization can threaten the executive’s office.

\textsuperscript{16} The global human rights polity includes international organizations, INGOs, states concerned with human rights, companies concerned with human rights.
can increase the probability of being turned out of office by inducing more mobilization because fewer public goods can be offered. Lastly, increased mobilization increases the chance of legal sanctions (see last section) as domestic groups often use legal channels as a strategy to attain goals. What follows are ways in which NHRIs increase mobilization.

In order for individuals to mobilize, they must overcome collective action problems (e.g. Olson, 1971). Torture can lower the probability victims overcome this hurdle because it can cause (perceived) isolation, fear of the government, and paranoia (e.g. Scarry, 1985; Sussman, 2005). Just as NGOs increase information about rights and what actions individuals can pursue when those rights are abused (Smith-Cannoy, 2012), NHRIs release information about their existence and the role they play. When individuals come to the office to lodge complaints, they are informed of the state’s international obligations. Learning of other victims (either face-to-face at the NHRI or by information provided by them) also increases the probability individuals overcome collective action problems by decreasing feelings of isolation.

A skeptic may say that a traumatized, fearful victim would not visit an NHRI to lodge a complaint. If true, NHRIs still reach individuals with annual reports and events designed to raise awareness (e.g. Cardenas, 2012; Carver, 2012; Okafor, 2012). These outreach programs often include aggregated statistics of the types of complaints made against the state, thus helping victims feel less isolated and more likely to overcome collective action problems to mobilization.

Public outreach also informs non-victims. Often, people mobilize in support of others’ rights (e.g. Marx and Useem, 1971; Myers, 2008; Russo, 2014); they need not be tortured themselves to advocate for the abolition of torture. The release of public reports increases information about the conditions to citizens within a country. Citizens oppose torture, even when they believe it may thwart a terrorist attack (Gronke et al., 2010). Thus, if citizens believe their government is torturing, a subset will become upset enough to mobilize. Non-victims mobilizing can be just as dangerous to an executive’s power, if not more so, given larger numbers and more resources. Also, an increase in non-victim mobilization may increase victim mobilization by decreasing the collective action problem experienced due to isolation. It should also be noted that mobilized activists often work with and for NGOs. NGOs increase domestic mobilization by supplying resources and training to groups (Murdie and Bhasin, 2011) and lobbying international actors (Risse and Sikkink, 1999; Murdie and Davis, 2012), as mentioned above. Murdie and Davis
find that NGOs are most effective when NGO domestic presence is strong. NHRIs often work with NGOs, potentially strengthening the domestic environment, thus making the NGO actions (including mobilizing the public) more effective. NHRIs are important institutions for informing potential mobilizers about international commitments and how victims are not alone. Mobilization decreases the probability of the executive holding onto power, thus, in anticipation of the mobilization, an executive that makes commitments not to torture should torture less given NHRIs’ ability to provide information.

NHRIs are domestic institutions that provide information to domestic and international audiences about legal obligations and compliance that ultimately effects actors’ pursuit of legal redress and mobilization. Both the legalization and mobilization bring potential costs to the executive in the form of decreased rents and/or increased probability of losing power. In expectation of these costs, the executive should torture less often:

Hypothesis: States already party to the CAT will torture less often when they have an NHRI.

1.4 Data and Methods

1.4.1 Dependent Variable

The outcome of interest is state torture practices, which I operationalize using an ordinal scale ranging from 0 to 2 taken from the Cingranelli-Richards Human Rights Dataset (CIRI) (Cingranelli and Richards, 2012). The measure maps consistently with my conception of torture. According to the CIRI database, “[t]orture refers to the purposeful inflicting of extreme pain, whether mental or physical, by government officials or by private individuals at the instigation of government officials. Torture includes the use of physical and other force by police and prison guards that is cruel, inhuman, or degrading. This also includes deaths in custody due to negligence by government officials. A score of 0 indicates that torture was practiced frequently in a given year; a score of 1 indicates that torture was practiced occasionally; and a score of 2 indicates that torture did not occur in a given year” (Cingranelli, Richards and Clay, 2014, 3). Coders used content analysis on Amnesty International reports supplemented by U.S. State reports to create the measure.

See Murdie (2014) for an interesting argument suggesting NGOs should not be treated as monolithic organizations, but rather can be grouped according to their motivations.
1.4.2 Explanatory Variable

The independent variable used to test the hypothesis is an interaction term that is the product of whether a country has ratified, acceded, or succeeded (hereon referred to as ratified) to the CAT and has an NHRI. The year a country ratified the CAT was coded a 1, as was every year thereafter; as opposed to 0 for those years in which the country had not. I obtained the data for CAT ratification from the United Nations treaty website (UN, 2012). Of the 3778 country years, 1800 had ratified, acceded, or succeeded to the CAT.

Whether a country has an NHRI is a dichotomous measure, coded 1 if an NHRI exists in a given country-year and 0 otherwise. I collected this data by consulting the International Coordinating Committee of National Institutions for the Promotion and Protection of Human Rights’ (ICC) global list of NHRIIs (ICC, 2012) to determine the official NHRI in a given country and then using Koo and Ramirez (2009, Appendix 3) to determine the year the state established the NHRI. The year a country establishes an NHRI and each subsequent year is coded 1 as opposed to 0 (when the NHRI did not exist). Of the 3778 country years, 1333 had an established NHRI.

1.4.3 Control Variables

Democracies tend to commit rights violations less frequently, including the violation of torture (Conrad and Moore, 2010; Davenport, 1995, 1999; Davenport and Armstrong, 2004). I measure democracy using the Polity IV database (Marshall, Gurr and Jaggers, 2012) which assigns countries a score from -10 to 10 (from most autocratic to most democratic) based on six component measures that record key qualities of executive recruitment, constraints on executive authority, and political competition. I operationalize democracy by creating a dichotomous variable that takes the value 1 if the Polity score is greater than or equal to 6, and 0 otherwise. I expect democracies to torture less often.

Often referred to as the “more murder in the middle” phenomenon, transitioning democracies and anocracies abuse citizens more often (Fein, 1995; King, 1998; Regan and Henderson, 2002).

---

18 The Polity IV researchers suggest democracies defined at Polity between 6 and 10, anocracies between -5 and 5, and autocracies at -10 to -6 (see http://www.systemicpeace.org/polityproject.html. I thank an anonymous reviewer for the suggestion of dummies for democracy and anocracy. I have also used the Polity scale for democracy and a dummy variable for transitioning states with almost identical results, but the current specification makes for a more logical model fit.

19 Recall this would result in an increase in the dependent variable as it measures protection of rights.
In order to control for this particular institutional arrangement, I again use Polity to create a dichotomous variable that takes the value 1 when the Polity score is between -5 and 5, and 0 otherwise.\textsuperscript{20}

I control for population because a large population may lead to more human rights abuses. Large population increases the occurrence of rebellion (real or perceived), thus a leader willing to violate human rights will do so more often (Poe, Tate and Keith, 1999). Larger populations also place stress on resources (Henderson, 1991), leading to more demands from citizens to their government that the government may not be able to provide. I take the log of the population that I acquired from the World Bank database (World Bank, 2013). I expect that as the log of the

\textsuperscript{20}Autocracy (Polity -10 to -6) is the referent category and is not included in the model.
population increases, the torture score will decrease (more torture) when controlling for all other factors.

Lower economic standing leads to higher human rights abuses (Davenport, 2007a; Poe, Tate and Keith, 1999). Poorer countries have less money to redistribute leading to higher citizen grievances that could be quelled with coercive acts e.g. torture. I control for a country’s level of economic development by including its yearly energy consumption (kg oil consumed per capita). I expect that as the log of energy consumption increases, torture will decrease, resulting in higher CIRI torture scores, on average.

I control for the presence of an independent judiciary because previous studies find an association with freer societies with fewer human rights abuses (Cross, 1999; Keith, 2002; LaPorta et al., 2004). I use the latent variable created by Linzer and Staton (2011) based on eight commonly used proxies and indicators for judicial independence, as it “makes use of the general agreement among the indicators, yet addresses concerns” with these same indicators (13). It is a continuous variable between 0 and 1, with higher values representing more judicial independence. I expect that as the independent judiciary score increases, a country’s CIRI torture score will increase (less torture).

I include in the model whether a country has an independent judiciary and has ratified the CAT. As discussed above, the domestic judiciary will be more effective when acting in accordance with an agreed upon treaty (e.g. Lupu, 2013; Powell and Staton, 2009).

Transitioning democracies torture less often if they commit to the Convention against Torture (Simmons, 2009). In order to control for this relationship I interact the transition variable with CAT ratification.

---

21 Energy Consumption is a more reliable and available measure than GDP, and the two are highly correlated capturing the same concept (Jackman, 1973).

22 Clague et al. (1999); Feld and Voigt (2003); Howard and Carey (2004); Tate and Keith (2009); Cingranelli and Richards (2010a); Marshall and Jaggers (2010); Rios-Figueroa and Staton (2010)
Finally, I follow convention and control for autocorrelation by including a one-year lag of my dependent variable (Beck and Katz, 1995).

### 1.4.4 Statistical Model

The data are an unbalanced time-series cross-section (TSCS) covering 153 countries from the years 1981 to 2007. The unit of analysis is the country-year. Due to the ordinal nature of the dependent variable, I fit an ordered logit. When dealing with TSCS data, one must always be wary of autocorrelation and heteroskedasticity (Poe and Tate, 1994). Heteroskedasticity can lead to inaccurate standard errors, while autocorrelation can lead to both inaccurate standard errors and biased coefficients (Stimson, 1985). Beck and Katz (1995) suggest panel-corrected standard errors when using TSCS data, but this is probably a poor method when N is much larger than T (Hoechle, 2007, 284). For this reason, I use clustered standard errors on country to account for heteroskedasticity. As mentioned above, I include a lagged dependent variable on the right hand side of the equation to account for autocorrelation (Beck and Katz, 1995).

### 1.5 Results

All coefficient estimates are given in Table 1.1. The coefficient for the explanatory variable (the interaction of CAT ratification and Existence of an NHRI) is positive and significant at standard levels.

Figure 1.5 presents plots of the explanatory variable and its constituent terms for the hypothesis test. Consistent with past work, ratifying the CAT is associated with more state torture (e.g. Hathaway, 2002; Hill, 2010; Hollyer and Rosendorff, 2011). However, when states ratify the CAT and have an NHRI, state torture decreases. The coefficient on the interaction term is signed in the expected direction and significant at conventional levels ($\beta = 0.435$, $p = 0.046$).

Because the model is non-linear in outcome probabilities and the independent variable is binary

---

23 Three of the variables do not meet the assumption of parallel slopes (CAT ratification, International War, and Lagged Torture), but the overall model meets the assumption. Fitting a partial proportional odds model does not substantively change the beta coefficient (0.41) or standard error (0.27) on the explanatory variable or its constituents in a substantial way; therefore, the results are the same and not model dependent.

24 Another way to deal with autocorrelation is use of the Prais-Winston transformation (Plumper, Troeger and Manow, 2005). Estimating the Prais-Winston does not require a lagged dependent variable (which absorbs much of the time-series variance), so statistical results are “easier” to find, making my model choice a more conservative test. Results are similar; beta coefficient: 0.14, SE: 0.07.

25 I used STATA 12 for statistical analyses.
Table 1.1: Effectiveness Ordered Logit Results, 1981–2007

<table>
<thead>
<tr>
<th></th>
<th>Coefficient</th>
<th>(Std. Err.)</th>
</tr>
</thead>
<tbody>
<tr>
<td>CAT and NHRI</td>
<td>0.435</td>
<td>(0.218)</td>
</tr>
<tr>
<td>CAT</td>
<td>-0.821</td>
<td>(0.227)</td>
</tr>
<tr>
<td>NHRI</td>
<td>-0.214</td>
<td>(0.190)</td>
</tr>
<tr>
<td>Democracy</td>
<td>-0.739</td>
<td>(0.229)</td>
</tr>
<tr>
<td>Transition</td>
<td>-0.649</td>
<td>(0.264)</td>
</tr>
<tr>
<td>Civil War</td>
<td>-0.890</td>
<td>(0.197)</td>
</tr>
<tr>
<td>Internat War</td>
<td>-0.059</td>
<td>(0.241)</td>
</tr>
<tr>
<td>Econ Standing</td>
<td>0.121</td>
<td>(0.062)</td>
</tr>
<tr>
<td>Population</td>
<td>-0.259</td>
<td>(0.047)</td>
</tr>
<tr>
<td>Jud Ind</td>
<td>3.346</td>
<td>(0.557)</td>
</tr>
<tr>
<td>CAT and JudInd</td>
<td>0.281</td>
<td>(0.365)</td>
</tr>
<tr>
<td>CAT and Trans</td>
<td>0.005</td>
<td>(0.325)</td>
</tr>
<tr>
<td>Torture(_{r-1})</td>
<td>1.903</td>
<td>(0.105)</td>
</tr>
<tr>
<td>Cut1</td>
<td>-1.459</td>
<td>(0.835)</td>
</tr>
<tr>
<td>Cut2</td>
<td>2.036</td>
<td>(0.843)</td>
</tr>
</tbody>
</table>

n = 2931

Figure 1.2: Effect of Interaction Term and Constituents (with 95% Confidence Intervals)
Brambor, Clark and Golder, 2006; King, Tomz and Wittenberg, 2000; Long, 1997), I present the first difference in Table 1.2. Using Clarify (King, Tomz and Wittenberg, 2000; Tomz, Wittenberg and King, 2003), I obtain the first difference that an “average” country will fall into one of the three categories of the dependent variable when a country has ratified the CAT and goes from no NHRI to having an NHRI. Most importantly, countries substantively find themselves in the most egregious category (0) five percent less often when adopting an NHRI. Adopting an NHRI allows states that have ratified the CAT to move out of a situation where torture is routine (0) to situations where only “a few instances” occur (Cingranelli and Richards, 2010b, 17). The increase in states in category 1 is promising as well as more states are moving from 0 to 1, than are moving from 2 to 1 evidenced by the negative sign on 0 and positive sign on 2.

Table 1.2: Probability of Changing Torture Categories for CAT Ratifier when It Adopts an NHRI

<table>
<thead>
<tr>
<th>Torture Category</th>
<th>Probability</th>
<th>Std. Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>-0.049</td>
<td>0.034</td>
</tr>
<tr>
<td>1</td>
<td>0.036</td>
<td>0.026</td>
</tr>
<tr>
<td>2</td>
<td>0.013</td>
<td>0.009</td>
</tr>
</tbody>
</table>

Many of the control variables perform as expected (e.g. transitioning/anocratic states, civil war, economic standing, population, judicial independence). However, a few variables warrant further discussion. The presence of an NHRI is negatively signed and fails to meet standard statistical significance suggesting that in the absence of CAT ratification the impact of NHRIs on state torture practices is nil. The result is counter to Cole and Ramirez (2013), who find NHRIs independently lead to better physical integrity rights protection (including torture as well as extra-judicial killings, political imprisonment, and disappearances). Cole and Ramirez (2013, 7)

---

26 I set dichotomous variables to their medians and continuous variables to to their means. The result is an average country that is anocratic, is not at war (civil or international), has an economic standing on par with Brazil in 2006, a population of around 11 million, and judiciary with an independence score of 0.452.

27 This is the appropriate first difference to test my hypothesis, however in the full disclosure the first difference that takes a country with an established NHRI and changes the state from a CAT non-ratifier to ratifier yields opposite results. Not only does the first difference presented answer my question, but it is more appropriate given that most states ratify the CAT before they establish an NHRI (modal CAT ratification year is 1987, modal NHRI adoption year is 1996). The disparate first difference does present a puzzle about timing of institutional adoption worth exploring in future work.

28 Throughout I assume movement from adjacent categories. It is possible that countries go from 0 to 2, and vice versa, but unlikely.

29 Judicial independence has the highest substantive effect on torture practices in this model. Despite that fact, knowing about NHRI effectiveness for states that commit to international agreements is worthwhile.
suggest NHRI effectiveness with respect to physical integrity rights is due to the “principled inviolability of physical integrity rights or widespread cultural opposition to civil and political rights.” As mentioned above, freedom from torture is one of, if not the most, sacred of the international human rights. The disparate results may be due to different mechanisms involved in each of the physical integrity rights violations and how NHRI.s interact with governments vis-a-vis these rights. For instance, the decision to torture is probably affected by different considerations than whether to disappear someone. These results suggest Davenport’s (2007a) call to study disaggregated human rights rather than a collection of them will lead to further insight on institutional capacity to restrain. The disparate results warrant further investigation on NHRI effectiveness.

The democracy control is signed opposite of expectations and reaches conventional statistical significance suggesting democracies torture more often compared to the baseline category (autocracies). Although surprising given past scholarship has found such a robust correlation between democracy and respect for rights that Davenport (2007a) refers to the “Domestic Democratic Peace,” it is important to remember that democracy is a system of governance composed of different institutions and norms (Conrad and Moore, 2010; Conrad, Hill and Moore, 2014; Li, 2005; Souva, Smith and Rowan, 2008). The CAT and NHRI.s represent liberal legal institutions associated with the principles of democracy. However, elections also define democracies. Conrad, Hill and Moore (2014) show elections are associated with more torture because they are fundamentally majoritarian, and often leaders target minorities or “others” (Conrad, Hill and Moore, 2014; Conroy, 2000; Rejali, 2007).

The sign on the interaction of CAT ratification and judicial independence fails to reach conventional statistical significance (although it is signed in the expected direction), counter to (Powell and Staton, 2009). It is unclear to this researcher why this is. Further exploration of NHRI and judicial interaction may bear more information on this unexpected result.

The coefficient on the interaction of CAT ratification and transitional democracies also fails to reach conventional statistical significance (though the sign is in the predicted direction). When including NHRI.s in the model, Simmons’ (2009) result does not hold. The result makes sense given that many transitioning states established NHRI.s in an effort to become legitimate members of the international community (Reif, 2000).
Lastly, some may reasonably argue that the NHRI variable is not completely exogenous in the model. For this reason, I also estimated the ordered logit with every independent variable lagged by one year.\textsuperscript{30} The coefficient on the explanatory variable remains positive (0.387) with a slightly increased p-value (0.077).

### 1.6 Conclusion

The present contributes to the growing research agenda on the effectiveness of international law conditional on domestic conditions (e.g. Dai, 2014). For international torture law, NHRI\textsubscript{s} are an important liberal institution in the domestic political environment alongside effective judiciaries, (e.g. Goodman and Pegram, 2012; Lupu, 2013; Powell and Staton, 2009) and other domestic conditions (e.g Simmons, 2009). NHRI\textsubscript{s} offer information to interested actors that increase the chance they mobilize and pursue legal redress, thus causing an executive to decrease her expected utility to torture.

This study also contributes to research on disaggregated outcomes and explanations. Institutions that provide information are vital for making international law, such as the CAT, effective for the secretive practice of torture. Also, examining liberal legal institutions apart from democracy helps isolate which parts of democracy (in this case a liberal institution) contribute to the “domestic democratic peace” (Davenport, 2007\textsuperscript{a}).

The point of investing to make NHRI\textsubscript{s} more effective suggests further research into NHRI heterogeneity. The present study follows other quantitative pieces (e.g Kim, 2013; Koo and Ramirez, 2009)\textsuperscript{31} in treating NHRI\textsubscript{s} as dichotomous (i.e. absent or present) although they differ with respect to budgets, rights protected, allowed actions, and composition (e.g Goodman and Pegram, 2012; Mertus, 2009). Some NHRI\textsubscript{s} may be better at holding leaders to their word when they ratify the CAT. Which states adopt which kinds of NHRI\textsubscript{s}? What actors influence the adoption and design of different NHRI\textsubscript{s}? By seeking answers to these questions, future research should tease

\textsuperscript{30}Some may suggest a matching algorithm to preprocess the data in order to simulate experimental research design (e.g. Simmons and Hopkins, 2005; Hill, 2010). However, this method is unusually challenging given the explanatory variable of interest is an interaction term of two dichotomous constituent terms in the model. I leave this challenge to future research.

\textsuperscript{31}Cole and Ramirez (2013) begin to disaggregate by broadly grouping NHRI\textsubscript{s} based on their organizational structure and mandate (classic ombudsman, human rights commissions, and human rights ombudsman).
apart the differences in NHRI\textquotesingle s ability to offer information and keep governments accountable to their international commitments, such as the CAT.

Further research should also look at NHRI effectiveness with respect to other human rights treaties. Cole and Ramirez (2013) contend NHRI\textquotesingle s focus on physical integrity rights. Would NHRI\textquotesingle s effect the Convention on the Rights of Persons with Disabilities or the Convention on the Elimination of Discrimination against Women? If so, what mechanisms would be responsible? How does NHRI design effect these mechanisms? This study shows NHRI\textquotesingle s responsible for making the CAT effective by increasing information. Further research will show if NHRI\textquotesingle s are effective for the rest of the international human rights legal regime, and how.

Future scholarship should explore the timing of institutional adoption. For instance, this paper tested the effects of establishing an NHRI after the state has ratified the CAT. This institutional adoption timing is most appropriate given my research question and the fact that most states ratify the CAT before they adopt an NHRI (modal CAT ratification year, 1987; modal NHRI adoption year, 1996). However, performing a similar first difference, but switching the timing of adoption (NHRI first, then CAT), suggests an increase in the probability that states are abusive. Future research should explore this counter–intuitive result exploring when and why states would adopt different human rights institutions and the repercussions of these adoption timing decisions.

The findings in this paper suggest important policy implications. Hafner-Burton (2013) prescribes a triage strategy to make “human rights a reality.” Instead of an all–inclusive international legal regime, those concerned with human rights practices should devote their limited resources where the probability for success is highest. The current international human rights treaty regime is built on a model of universal inclusion in order to change norms throughout the system (Finnemore and Sikkink, 1998; Hafner-Burton, 2013). But, taking domestic conditions and availability of resources seriously, states and organizations should invest in creating new NHRI\textquotesingle s where they do not exist.\textsuperscript{32} Given that a top concern for NHRI effectiveness is budget constraints (e.g. Renshaw and Fitzpatrick, 2012; Smith, 2006), investing in existing NHRI\textquotesingle s to make them more capable of making international treaty obligations meaningful would be an effective strategy to improving respect for rights internationally. There is reason to be optimistic, as is evi-

\textsuperscript{32}If not after investing in steps to make the judiciary independent given its large substantive effects in the model.
denced by the requirement of a national preventative mechanism (often designated as the NHRI) in the Optional Protocol to the Convention against Torture.
CHAPTER 2

THE ROLE OF REPUTATION IN THE NORMS CASCADE: THE CASE OF NHRI ADOPTION

Certain norms (whether they be ideas or behavior) begin cascading after reaching a tipping point (Finnemore and Sikkink, 1998). What explains the tendency for states to adopt certain norms? When do individual states choose to be a part of the process that seems so inevitable to the outside observer in retrospect? The questions are particularly salient with respect to the international human rights regime, as the regime is built upon the commitment states make to a number of institutions including treaties and organizations. Why do states allow themselves to be caught up in a cascade by adopting institutions that can potentially encroach on their sovereignty? While scholars have studied the reasons for state commitment to international human rights institutions (e.g. Mansfield and Pevehouse, 2006; Hathaway, 2007; Vreeland, 2008; Simmons, 2009), this paper seeks to explain the importance of international reputation to the process of the cascade by studying a specific domestic institution tied to the international human rights regime: national human rights institutions (NHRIs).

One cannot deny the speed at which states have adopted NHRIs – domestic institutions charged with promotion and protection of human rights – over the past few decades. The first institutions concerned with government accountability (Ombudsmen), and thus the predecessors of NHRIs as we know them today, date back to the 19th and early 20th centuries.\(^1\) From 1990 to 1993 the total number of NHRIs almost doubled from 22 to 41. Presently the International Coordinating Committee (ICC) of National Institutions for the Promotion and Protection of Human Rights\(^2\) lists 147 NHRIs, globally.\(^3\) Some scholars studying this rapid cascade of adoption focus on specific types of institutions, such as the UN (Cardenas, 2003) or NGOs (Kim, 2013). Others

\(^1\) Sweden (1809); Finland (1919)

\(^2\) The ICC is an international association created by NHRIs for NHRIs. It seeks to promote and strengthen NHRIs by engaging and promoting within the UN, encouraging cooperation among NHRIs, and accrediting NHRIs, among other tasks. See http://nhri.ohchr.org/EN/Pages/default.aspx for more details.

\(^3\) This number exceeds the number of states in the world due to the existence of NHRIs in some non-states (e.g. Palestine, Northern Ireland) and multiple NHRIs in some states (e.g. Sweden, Switzerland, Tajikistan).
seek to more generally describe the diffusion of the NHRIs with rationalist (Koo and Ramirez, 2009) and sociological/cultural (Koo and Ramirez, 2009; Pegram, 2010) arguments.

I contribute to these more general explanations by offering the argument that states seek to adopt norms such as NHRIs in order to become more socialized, but only when in their rationalist interests do so – states act both socially and rationally simultaneously. I argue states care about their socialization into the club of states, thus they care that other states see them as such. In short, states care more or less about their international reputation. Adopting an NHRI can signal to other states that the state in question wishes to be associated with the club. But a state will only do so when the marginal benefits outweigh the costs. Holding costs constant,4 a state can estimate its current reputational standing among international actors by observing the shaming behavior directed at them. I argue that the affect of international shaming’s effect on NHRI adoption follows an inverted-U shape. Those states that are almost never shamed or shamed all of the time won’t find it marginally beneficial to adopt an NHRI. As state shaming increases, the probability states adopt an NHRI increase to an extent at which point it drops again.

I test the proposed hypothesis by estimating a Cox proportional hazards model on a global sample from the years 1991-2009 and find support for the argument.

The current project lends understanding to the broader International Relations literature on norms cascades (Finnemore and Sikkink, 1998). Specifically, I show that international reputation is an important mechanism for whether or not a norm cascades after reaching its tipping point. More specifically, the effect of international shaming on state behavior follows an inverted–U pattern rather than a linear one. The project also has implications for the human rights literature by pointing the importance of reputation in the decision of states to adopt institutions with the power to limit their precious–guarded sovereignty. Finally, I join others (Cardenas, 2003; Koo and Ramirez, 2009; Pegram, 2010; Kim, 2013) in explaining the rapid diffusion of the NHRI model throughout the world.

The rest of the paper proceeds as follows. First I briefly introduce the reader to the institution under study: NHRIs. Next, building on work by others, I propose an argument for the impor-

4 Future work should relax the assumption of constant costs
rance of state reputation in NHRI adoption. Then I test the proposed hypothesis, present the results, and discuss their importance. Lastly I conclude.

### 2.1 The Liberal Order and NHRI\textsc{s}

The extent to which a state adheres to human rights norms helps other states make value judgements about their willingness and ability to be a part of the current international order. In effect, states have reputations.

After WWII, the Allies pushed the formation of the United Nations (Korey, 1998). The world was war-weary and negotiated an organization that would help to overcome collective action problems associated with cooperation that could lead to peace and stability. Due to the diligence of a group of American NGOs\textsuperscript{5}, the Preamble of the UN Charter stresses “to reaffirm faith in fundamental human rights, in the dignity and worth of the human person, in the equal rights of men and women and of nations large and small” which is tied inextricably to the ends of “promotion of the economic and social advancement of all peoples.” States believed these actions could “save succeeding generations from the scourge of war.” The acceptance of the tenets of the United Nations mark a change in appropriate international behavior for states.

In order to legitimate oneself in the international order, states must recognize and heed the importance of the liberal order—including human rights\textsuperscript{6}. For a state to signal concern with human rights, it must take actions that other states can observe. These behaviors allow states to create beliefs about other states—in essence, in the international society, states have reputations.

Commitment to international institutions (e.g. treaties) became important observable behavior. The Universal Declaration of Human Rights and the first international human rights treaty in 1948\textsuperscript{7} led to a preponderance of later human rights treaties.\textsuperscript{8}

However, given the almost immediate increase in the importance of international institutions after the establishment of the UN, one has a hard time observing a clear tipping point that

\textsuperscript{5} A group of American NGOs is responsible for the extensive use of human rights language in the UN Charter distinguishing “it most sharply from the Covenant of the League of Nations”(John P. Humphrey, quoted in Korey, 1998, 30).

\textsuperscript{6} Other normative liberal concerns made important by the UN include settling conflict peacefully, acceptance of international legal principles, democratization, and trade liberalization (UN Charter Article 1).

\textsuperscript{7} The Convention on the Prevention and Punishment of the Crime of Genocide

\textsuperscript{8} To date, nine core UN human rights treaties exist, with dozens more international instruments (of the High Commissioner for Human Rights, 2006; Bayefsky, 2014).
leads to a cascade. A clear tipping point allows the observation of the norm cascade, which allows the exploration of the mechanism of the cascade. Focusing on NHRI adoption allows us a much clearer view of the norm cascade. Despite adoption by a few entrepreneurs (e.g. Sweden, Norway, Denmark, New Zealand) and a couple international calls for domestic institutions to make human rights standards a reality on the ground, the Paris Principles – a document that defined NHRI s and promoted their importance in the international human rights regime – mark the tipping point for NHRI adoption.

### 2.2 NHRI Adoption

Ombudsman offices represent the first institutions now considered NHRIs. Although ombudsman offices do not follow explicit human rights mandates, they do perform tasks in order to discover and correct government misconduct and maladministration. The ombudsman model of government accountability, with its roots in 19th century Sweden, spread throughout Scandinavia (Finland, 1919; Denmark, 1955; Norway, 1962) to other parts of the world (New Zealand, 1962; Australia, 1977) (Cheng, 1968; Rowat, 1973; Koo and Ramirez, 2009).

Meanwhile, the United Nations and the international community placed greater emphasis on the protection of individual rights. Some saw domestic institutions as able to supplement the work of the UN Commission on Human Rights and the increasing international human rights treaty regime (Cardenas, 2003). The UN’s hard work got NHRI s as a norm to the tipping point.

In 1978, the Commission on Human Rights organized a seminar to more clearly define a “national and local institutions for the promotion and protection of human rights” opening the way for states to adopt NHRI s as norm entrepreneurs. The seminar highlighted a political consensus that domestic processes may enhance the international human rights regime. A handful of (mostly Western) countries adopted domestic institutions focused on human rights.\(^9\)

The International Workshop on National Institutions for the Promotion and Protection of Human Rights (Paris, 1991) represents the tipping point for NHRI adoption. Participants (includ-

---

9 In 1946 and again in 1978, the UN highlighted the importance of domestic institutionalization for human rights outcomes, but the idea of NHRI s as a specific type of institution remained vague and hardly caught on (Cardenas, 2003; Mertus, 2009).

10 Although many of these states fall easily into what (Simmons, 2009) refers to as “sincere” states, a number of more notable exceptions occur (e.g. Benin, Gabon, Zambia). Future research should analyze what makes norm entrepreneurs willing to be such early adopters.
ing NHRI, States, UN agencies, IGOs, and NGOs) drafted a document specifying NHRI role, composition, status, and function; essentially defining them by documenting clear standards of what make an institution an NHRI.\textsuperscript{11} The document, now known as the Paris Principles, quickly gained a UN Commission on Human Rights endorsement in 1992\textsuperscript{12} and adoption by the General Assembly in 1993\textsuperscript{13} (UN, 1993).

The cascade began. States started adopting NHRI at breakneck speeds. Whereas only one state adopted an NHRI in 1990 (Namibia), 18 more states adopted from 1991-1993 alone. As the 1990s continued, more and more states adopted NHRI. (Cardenas, 2003) does an excellent job of highlighting the importance of the UN to get the concept of the NHRI to the tipping point. But what explains when states decide to adopt the institution within the norms cascade?

\section*{2.3 Past Explainations}

After 1991, NHRI quickly diffused throughout the global system. (Koo and Ramirez, 2009) pitted “Neo–Realism” vs. “World Polity” finding mixed support for each depending on the institution studied,\textsuperscript{14} with more support for the World Polity hypotheses – which amount to correlations between numbers of international institutions or human rights institutions and NHRI adoption. They tested their hypotheses on a temporal span including the entrepreneur and cascade step of the norm life cycle (1966-2004), but their results hint at a couple underlying lessons. States consider both their immediate self–interest and the opinions of others when deciding to adopt an NHRI – a theme I return to below as I lay out my theoretical argument.

In explaining the rise in NHRI adoption, (Pegram, 2010) also identified a diffusion process. Drawing on Goodman and Jinks (2004), he proposes three possible mechanisms for diffusion: coercion, acculturation, and persuasion.\textsuperscript{15} Diffusion refers to a process of “uncoordinated inter-
dependence” (Elkins and Simmons, 2005, 33) that results in a given state’s probability of adoption increasing as the number of previous adopters increases (Strang, 1991). With respect to NHRIIs, we know empirically that the process occurs: as the number of NHRIIs increase, the probability a state adopts an NHRI increases (Koo and Ramirez, 2009). But how do we know when states will decide to adopt an NHRI? What drives the finding that corroborates the process?

2.4 Diffusion, Reputation, and Status

Diffusion processes often lack an underlying theoretical process that yields testable implications. (Kim, 2013) fills some of the gap suggesting human rights INGOs (also called human rights organizations (HROs) (Murdie and Bhasin, 2011)) drive developing state adoption of NHRIIs by one of two theoretical pathways – the mediation of discourse between states and the UN and the mobilization of international shame.

The first of these pathways echoes the importance of the UN’s agenda and standard setting through international conferences and resolutions (Cardenas, 2003), which serve to strengthen the norm of NHRIIs as desirable thus strengthening states’ beliefs about the potential benefits of NHRI adoption. The second hints at those benefits. The mobilization of shame implies that a socially accepted behavior exists. Adhering to the socially accepted behavior, adopting an NHRI, states may gain reputational benefits leading to international status.

State status and reputation underly, often implicitly, much of the literature about NHRI diffusion. Status refers to “collective beliefs about a given state’s ranking on valued attributes” (Larson, Paul and Wohlford, 2014, 7). State status is a second-order belief – a belief that a group of observers holds some belief (Dafoe, Renshon and Huth, 2014). Reputation, on the other hand, represents a first–order belief: “the beliefs one actor holds about another actor’s characteristics or behavioral tendencies (Dafoe, Renshon and Huth, 2014, 374). First and second–order beliefs need not agree at any given time, but they tend to correlate highly with each other, as first–order beliefs inform second–order beliefs (Dafoe, Renshon and Huth, 2014, 374). In this case, the reputation of a state informs that state’s status. Status accrues benefits to states proximately and distally. Proximately, theoretical and empirical evidence from economics and social psychology presented norm. This logic leads to questions of downstream outcomes, such as design, once a state adopts an NHRI. Here I focus on diffusion by acculturation, and refer to it as diffusion.
suggest individuals seek status as an end unto itself (Bakshi and Chen, 1996; Huberman, Loch and Onculer, 2004). Distally, the collection of status as a good leads to (potential) diffuse material benefits in the future through at least two pathways. First, the higher the state’s status the higher the probability other states defer to that state (Lake, 2009), allowing that state to inform the future rules of the game. Secondly, those states higher in status increase their access to material rewards from others (e.g. access to markets). For instance, past research shows better human rights practices lead to more trade (Hathaway, 2005), aid (Cingranelli and Pasquerello, 1985; Poe, 1992; Lebovic and Voeten, 2009; Nielsen, 2013), and FDI Blanton and Blanton (2007).

States can adjust their status by increasing their reputations. By ascribing to a norm that shows promise for reputational benefit (i.e. a norm past the tipping point), states send signals to other states of their acceptance of the prevailing regime and their desire to play a part in the regime. As states adjust their beliefs about the adopting state, its reputation increases, thus increasing its social status, which in turn increases that state’s access to the benefits associated with status. Adopting an NHRI can lead to higher status. As the South Asian Human Rights Documentation Service put it: “If in the 1950s, the status symbol of a developing country was a steel mill, in the 1990s, apparently, it was a human rights commission” (Burdekin, 2007, quoted in Mertus (2009, 4)).

Not all states adopt NHRI because states weigh those benefits against potential costs and decide whether to adopt an NHRI, thus perpetuating the norm cascade, or not. Costs of adopting an NHRI include the direct material costs of establishing a new institution, as well as the sovereignty costs associated with a watchdog institution. As NHRI began to gain momentum as a normative institutional goal, states connected them to these ideals making it clear that reputational benefits existed for adoption. When benefits outweight costs, states adopt.

### 2.5 Importance of Shaming

Although we cannot observe reputation and status directly, public shaming represents a reasonable proxy for what states think of each other. A number of actors shame states including INGOs, the UN, and the media. They raise awareness about state violations of the liberal regime

---

16 This manuscript focuses on the benefits of adopting, holding the costs of adoption constant. Future research should explore the variation in cost of adoption.
such as human rights abuses and undemocratic behavior. States understand the importance of their reputation for their status in the global order, and so they often perpetrate abuses as secretly as possible (e.g. Rejali, 2007). By shining a light on these abuses and other abuses that may not have been covered up, but still needed a broadcast in order to inform states, shaming actors wish to make violation of liberal norms reputationally costly.

States use the shaming behavior directed at them to understand others’ views of them in the global order. The more shaming a state experiences, the more likely it will take action to stave the bleeding. Adopting a norm with clear reputational benefits – a norm cascading throughout the global community – can be just the band–aid needed. As established above, NHRI adoption began cascading in the early 1990s. (Cardenas and Fibbert, 2005) present a number of anecdotal accounts of Middle Eastern and North African states creating NHRIIs to correct human rights reputations soiled by shaming or to prevent shaming in the future. (Kim, 2013) finds the presence of human rights NGOs increases the probability developing states adopt NHRIIs, due partly to increased pressure.

While the amount of shame received should influence the adoption rate of NHRIIs, I expect a non–linear, inverted–U relationship. States must maximize marginal returns of adoption. As shaming increases, the benefits of adoption increase to some point, at which they begin to decrease again. After all, if the international community seems bent on shaming a state (fairly or not), why pay the costs of adoption? The most shamed states – United States, China, Turkey, Indonesia, Israel, and Russia – show a variation in liberal institutions and behavior. Those relatively liberal states included may not anticipate high enough marginal gains due to continuing to draw the spotlight despite other liberal measures. Those more illiberal states may not anticipate high enough marginal gains as their actions may be seen as less credible.

Those states receiving the least amount of shame should not anticipate high marginal gains either. Other actors already condone and/or ignore their actions. What will adopting another liberal institution do for their status? Adopting an NHRI represents the acceptance of a cascading

---

17 Although see (Hollyer and Rosendorff, 2011) for an account of governments openly abusing people in order to assert authority over the domestic population.
18 Shaming can also induce other costs such as increasing the probability the opposition mobilizes (e.g. Risse-Kappen, Ropp and Sikkink, 1999; Murdie and Bhasin, 2011).
19 Of which only Indonesia has an NHRI; based on the measure described below.
norm that can act as a band-aid for reputational scrapes or cuts. A band-aid won’t help a major, gashing wound; and a state without an injury does not need a band-aid.

Hypothesis: As a state’s level of shaming increases, its probability to adopt an NHRI increases to a point, in which it will begin to decrease.

2.6 Research Design

2.6.1 Spatial-Temporal Domain

I evaluate the hypothesis discussed above over the global population of countries that had not adopted an NHRI before the Paris Principles, thus excluding the norm entrepreneurs, as the logic for their adoption may be different (and should be studied in the future). The temporal span of the study, then, represents the norm cascade (1991 to 2009). The data are structured as an unbalanced time-series cross-section, with country-year unit of analysis.

2.6.2 Model

Because my interest lies in the point at which a state adopts an NRHI, I estimate a Cox Proportional Hazards model. I opt to use the Cox proportional hazards model (Cox, 1972) because it does not require assumptions about functional form (unlike parametric models), but it is able to incorporate covariates (unlike Kaplan-Meier) (Lee and Go, 1997). Given a vector of covariates, \( x \), the hazard function for a country year, \( i \), at time, \( t \), is modeled by

\[
h_i(t, x_i) = h_0(t) \exp(x_iB)
\]

2.6.3 Variables

Dependent Variable. The dependent variable is dichotomous taking the value 1 if an NHRI is legally adopted in a particular country in a particular year, and 0 otherwise. I coded the adoption time by consulting various sources – NHRI annual reports, NHRI and government

---

20 The Paris Principles easily represent the beginning of the norm cascade. The next step of the norm life cycle is internalization in which the norm is taken for granted. NHRIs have yet to achieve such a status.

21 The covariates include the explanatory and control variables discussed below.

22 Once a state adopts an NHRI, the rest of the country years are absent from the dataset because hazard models model time to “failure” (“failure” being adoption of an NHRI in this instance). For example, Jamaica adopted the Office of the Public Defender in 1999. From 2000 to 2012, Jamaica no longer exists in the dataset.
websites, third-party websites, and academic pieces. Figure 2.1 presents the number of NHRIs adopted in each year of the study.

Figure 2.1: Global Adoption of NHRIs by year

**Explanatory Variable.** Testing the hypothesis requires a variable that represents the amount of shame a state receives in a year. Shaming can come from NGOs, the media, IGOs, and other states (Franklin, 2008; Hafner-Burton, 2008; Murdie and Davis, 2012; Ramos, Ron and Thoms, 2007). Each actor behaves strategically, with different reasons to shame different governments. The different shaming behavior of different actors has lead to disparate results in the literature with respect to shaming. In order to avoid arbitrarily choosing from which shaming data source to use, I create a Bayesian latent variable capturing the concept of shaming based on shaming data from a number of sources. I include the number of Amnesty International press releases and background reports (Ron, Ramos and Rodgers, 2005), the number of *Economist* and *Newsweek* articles discussing human rights abuses (Ron, Ramos and Rodgers, 2005), whether a country was targeted by the United Nations Commission on Human Rights (Lebovic and Voeten, 2009), and the number of shaming events by human rights NGOs (Murdie and Bhasin, 2011). Figures 2.2 and 2.3 show the fifty most and least shamed states according to the latent variable model with

---

23 For detailed discussion of the latent variable construction see Appendix.
the blue dots representing point estimates of the mean and the lines incorporating the min and max amount of shaming. I include the squared term of the latent variable due to the posited functional form. I expect the coefficient of the squared term to be negative and the coefficient of the constituent term to be positive.

Fifty Most Shamed States

<table>
<thead>
<tr>
<th>Country</th>
<th>Average State Shaming, 1991–2012</th>
</tr>
</thead>
<tbody>
<tr>
<td>United States</td>
<td></td>
</tr>
<tr>
<td>China</td>
<td></td>
</tr>
<tr>
<td>Turkey</td>
<td></td>
</tr>
<tr>
<td>Indonesia</td>
<td></td>
</tr>
<tr>
<td>Israel</td>
<td></td>
</tr>
<tr>
<td>Russia</td>
<td></td>
</tr>
<tr>
<td>United Kingdom</td>
<td></td>
</tr>
<tr>
<td>India</td>
<td></td>
</tr>
<tr>
<td>Colombia</td>
<td></td>
</tr>
<tr>
<td>Rwanda</td>
<td></td>
</tr>
<tr>
<td>Myanmar</td>
<td></td>
</tr>
<tr>
<td>Burundi</td>
<td></td>
</tr>
<tr>
<td>Pakistan</td>
<td></td>
</tr>
<tr>
<td>Peru</td>
<td></td>
</tr>
<tr>
<td>Guatemala</td>
<td></td>
</tr>
<tr>
<td>Bosnia and Herzegovina</td>
<td></td>
</tr>
<tr>
<td>Sri Lanka</td>
<td></td>
</tr>
<tr>
<td>Brazil</td>
<td></td>
</tr>
<tr>
<td>Mexico</td>
<td></td>
</tr>
<tr>
<td>Iran</td>
<td></td>
</tr>
<tr>
<td>Egypt</td>
<td></td>
</tr>
<tr>
<td>South Africa</td>
<td></td>
</tr>
<tr>
<td>Sierra Leone</td>
<td></td>
</tr>
<tr>
<td>Algeria</td>
<td></td>
</tr>
<tr>
<td>Romania</td>
<td></td>
</tr>
<tr>
<td>Cuba</td>
<td></td>
</tr>
<tr>
<td>Chile</td>
<td></td>
</tr>
<tr>
<td>Nigeria</td>
<td></td>
</tr>
<tr>
<td>Kenya</td>
<td></td>
</tr>
<tr>
<td>Afghanistan</td>
<td></td>
</tr>
<tr>
<td>Congo</td>
<td></td>
</tr>
<tr>
<td>Greece</td>
<td></td>
</tr>
<tr>
<td>Morocco</td>
<td></td>
</tr>
<tr>
<td>Saudi Arabia</td>
<td></td>
</tr>
<tr>
<td>Azerbaijan</td>
<td></td>
</tr>
<tr>
<td>Syria</td>
<td></td>
</tr>
<tr>
<td>Haiti</td>
<td></td>
</tr>
<tr>
<td>Sudan</td>
<td></td>
</tr>
<tr>
<td>Bulgaria</td>
<td></td>
</tr>
<tr>
<td>Cyprus</td>
<td></td>
</tr>
<tr>
<td>Uzbekistan</td>
<td></td>
</tr>
<tr>
<td>Spain</td>
<td></td>
</tr>
<tr>
<td>Belarus</td>
<td></td>
</tr>
<tr>
<td>Japan</td>
<td></td>
</tr>
<tr>
<td>Albania</td>
<td></td>
</tr>
<tr>
<td>Cambodia</td>
<td></td>
</tr>
<tr>
<td>Tunisia</td>
<td></td>
</tr>
<tr>
<td>Iraq</td>
<td></td>
</tr>
<tr>
<td>Maldives</td>
<td></td>
</tr>
<tr>
<td>Togo</td>
<td></td>
</tr>
</tbody>
</table>

Figure 2.2: Fifty Most Shamed States

Control Variables. Readers may wonder why I use shaming about illiberal state behavior instead of the actual behavior of the state. Shaming more closely fits the theoretical argument of the subjective nature of status – actual behavior matters only as much as other states care
Figure 2.3: Fifty Least Shamed States
about it. But in order to test for the subjective effects I include actual behavioral and institutional information as controls.

Table 2.1: Adoption Control Variables

<table>
<thead>
<tr>
<th>Variable</th>
<th>Mean</th>
<th>Std. Dev.</th>
<th>Minimum</th>
<th>Maximum</th>
</tr>
</thead>
<tbody>
<tr>
<td>Democracy</td>
<td>0.12</td>
<td>0.93</td>
<td>-2.02</td>
<td>2.26</td>
</tr>
<tr>
<td>Human Rights</td>
<td>0.65</td>
<td>1.51</td>
<td>-2.74</td>
<td>4.69</td>
</tr>
<tr>
<td>Material Dominance</td>
<td>-0.20</td>
<td>0.98</td>
<td>-2.80</td>
<td>2.80</td>
</tr>
<tr>
<td>Age</td>
<td>64.07</td>
<td>57.60</td>
<td>0</td>
<td>196</td>
</tr>
<tr>
<td>Judicial Independence</td>
<td>0.51</td>
<td>0.32</td>
<td>0.02</td>
<td>0.99</td>
</tr>
</tbody>
</table>

States desire other states to be democratic, and thus other states wish to be considered democratic. To further make the point a cursory survey of many relatively non–democratic states yields names such as the Democratic People’s Republic of Korea, People’s Republic of China, or the Democratic Republic of the Congo (Clark, Golder and Golder, 2013). As states democratize, they can adopt liberal institutions to signal to other states their intention. They can join international organizations (Mansfield and Pevehouse, 2006, 2008) and international/regional human rights institutions (Moravcsik, 2000). NHRI s represent yet another liberal institutional states can adopt to signal their adherence to liberal democratic norms. How democratic a state is affects the marginal benefit of adopting an NHRI to that state’s reputation. States that are very autocratic will not be able to send credible signals by adopting an NHRI, thus their marginal benefit would be low. These states must enact more costly policies as other states closely scrutinize. The most illiberal states may need to invest in elections that are then made transparent in order to prove to the world they wish to belong to the liberal order. Likewise, the most democratic states need not prove themselves and stand to gain fewer marginal benefits from adoption. The states with middling levels of democracy – those transitioning or democratizing – stand to gain the highest marginal benefit from adopting an NHRI. Past work highlights the correlation of NHRI adoption and democratizing states (e.g Cardenas, 2003; Koo and Ramirez, 2009; Pegram, 2010). I expect an inverse–U relationship, thus I add a squared term in addition to the measure of democracy.

Democracy is a latent concept, so I use the unified democracy score (UDS) (Pemstein, Meserve and Melton, 2010). The UDS uses a Bayesian latent variable that results in a “cumulative ap-
proach...for every country–year...[that] is at least as reliable as the most reliable component measure” (Pemstein, Meserve and Melton, 2010, 1).

State human rights behavior may be related to a state’s decision to adopt an NHRI and the main explanatory variable–level of shaming. Committing to public institutions may bring potential sovereignty costs as they limit how states can deal with dissent.\(^{24}\) For this reason, as state respect for human rights increase, the costs of adoption should decrease (e.g. Downs, Rocke and Barsoom, 1996; Simmons, 2009). (Koo and Ramirez, 2009) find greater human rights respect increases the probability states adopt a human rights commission or human rights ombudsman, but not a classical ombudsman. In addition to the possible relationship with adoption, not surprisingly, human rights behavior correlates with the amount of shaming a state receives (e.g. Risse-Kappen, Ropp and Sikkink, 1999; Franklin, 2008; Murdie and Davis, 2012; Hill, Moore and Mukherjee, 2013).

For these reasons I include human rights respect as a control using a latent variable measure from (Fariss, 2014) constructed using a dynamic modeling process with many of the most used measures used to measure human rights (e.g. PTS (Gibney and Dalton, 1996), CIRI (Cingranelli and Richards, 2012), ITT (Conrad, Haglund and Moore, 2013)).

Although I focus on the normative aspects of status, material dominance also contributes to state status (e.g. Gilpin, 1981; Morgenthau, 1985; Lake, 2009). Aside from their coercive abilities, materially powerful states yield important influence and contribute to what is socially acceptable behavior, which also influences state status. States most materially dominant in the international hierarchy will gain little in status from NHRI adoption. Material dominance includes military and economic tools that help states exert influence over other states’ sovereignty (Lake, 2009). The United States’ lack of an NHRI provides the perfect example of the importance of material dominance.

Likewise, those states with very low material dominance will achieve smaller marginal gains from the adoption of an NHRI. The increases in reputation just won’t yield a high enough status to pay the costs of adoption. Moving up a few rungs on the social ladder won’t allow Palau to offer serious input into the world order. In fact, these extremely small and weak states do not even concern themselves with much of the politics of the globe. Indeed, many readers will

\(^{24}\)Here I simplify the argument and by recognizing potential costs. Different NHRIIs may impose different expected costs based on how the government designs them once established.
recognize these states as those most often dropped from data sets due to their relative non-involvement. Examples of states with low dominance include San Marino, Tuvalu, Sao Tome and Principe, and Kiribati.

A state’s material dominance should predict its propensity to adopt an NHRI in an inverted-U pattern. Those states very dominant and very weak do not gain enough to pay the costs of adoption.

Dominance being a latent concept, I create a Bayesian latent variable using several indicators. I obtain a majority of the indicators from the constituent variables of the Composite Index of National Capability (Singer, 1988). I include all of the variables from the CINC scores—energy consumption, iron and steel production, military expenditure, military personnel, total population, and urban population. The CINC score captures many variables considered important as military resources. I also include a couple variables in order to capture economic resources with respect to global dominance. The first of these is total exports. Those states that provide exports to other states gain hierarchical leverage as they supply needed goods and are paid to do so. I obtain the export trade data from (Gleditsch, 2002). I also include state GDP obtained from (World Bank, 2013). I log every variable to account for skewness in the data.

Status and reputation depend on other states’ beliefs. Those beliefs rely on heuristic signals given by the state in question over time. A new state has a smaller history from which to send those signals, thus more noise around the estimate of their reputation. Establishing liberal institutions may help a new state shrink the variance around other states’ beliefs about their type. Thus newer states may wish to adopt an NHRI. For this reason, I include state age as a control. I create the variable by subtracting the year a state was founded from the current year in the database. I obtain the founded years from (Gleditsch and Ward, 2013).

Judicial independence may affect whether a state adopts an NHRI. The human rights responsibilities of NHRIs and domestic courts may overlap. These shared responsibilities could cause domestic judiciaries to welcome NHRIs to the domestic human rights scene as allies. Indeed, (Peruzzotti, 2012) recounts the River Basin Crisis of Matanza–Riachuelo in Argentina where the Defensoria del Pueblo (Argentina’s NHRI) and the Supreme Court complemented each other’s

25 The 50 most materially dominant and the 50 least materially dominant can be found in the Appendix.
26 I used population data from (World Bank, 2013) due to the larger temporal coverage.
efforts in 2003–2004. However, the redundancy of priorities may cause more competitive relationships. For instance, the Panamanian Defensor received resistance from the Supreme Court prior to its constitutional adoption in 2004 (Pegram, 2012). I remain agnostic as to whether judiciaries will view a new NHRI as complementary or competitive substitutes, but as the courts become more independent, they should be able to affect NHRI adoption towards their preferences.

Judicial independence is also related to some of the explanatory and control variables. The independence of the domestic courts affects economic growth and trade (which are included in the dominance variable) (Barro, 1997; Souva, Smith and Rowan, 2008). Independent judiciaries also cause states to abuse rights less often (Powell and Staton, 2009; Simmons, 2009; Lupu, 2013).

I control for the independence of the judiciary including a latent variable created by (Linzer and Staton, 2011) that leverages eight other judicial measures from past scholarship.

Figure 2.4: Hazard Function of Adopting an NHRI, 1991-2009

By including the non-parametric hazard function, I control for the World Polity variables proposed by (Koo and Ramirez, 2009). Allowing the data to define the probability a state will adopt an NHRI at baseline. This represents the hard-to-measure or hard-to-model aspects of the socialization pressures on a state by other states and can be seen in Figure 2.4. As time goes on, the probability a state adopts an NHRI increases, as we expect from the importance of
human rights in World Polity theory (Koo and Ramirez, 2009). In robustness checks I include World Polity variables in the model, but high collinearity\textsuperscript{27} causes difficulties drawing inference on the variables. For instance, when I include NHRI density – the number of global NHRI s present in a given year – in the model, the results remain virtually the same, but the maximum likelihood estimator is unable to estimate a standard error for the coefficient.

\section*{2.7 Results}

<table>
<thead>
<tr>
<th>Table 2.2: Effect of Shaming on NHRI Adoption, 1991–2009</th>
</tr>
</thead>
<tbody>
<tr>
<td>\begin{tabular}{lccc} \hline &amp; Hazard Ratio &amp; Robust Std. Error &amp; P-value \ \hline Shaming &amp; 1.79 &amp; 0.58 &amp; 0.072 \ Shaming$^2$ &amp; 0.75 &amp; 0.13 &amp; 0.083 \ Control Variables &amp; &amp; &amp; \ Democracy &amp; 1.96 &amp; 0.52 &amp; 0.012 \ Democracy$^2$ &amp; 0.63 &amp; 0.12 &amp; 0.017 \ Human Rights &amp; 1.05 &amp; 0.16 &amp; 0.745 \ Dominance &amp; 1.58 &amp; 0.30 &amp; 0.016 \ Dominance$^2$ &amp; 0.58 &amp; 0.10 &amp; 0.001 \ Judicial Independence &amp; 0.33 &amp; 0.30 &amp; 0.223 \ Age &amp; 1.00 &amp; 0.00 &amp; 0.180 \ \hline \end{tabular}</td>
</tr>
</tbody>
</table>

Table 2.2 presents the regression results from the main model. Coefficients are given as hazard ratios. A hazard ratio over 1 represents an increase in the probability a state adopts an NHRI (a positive coefficient), whereas a hazard ratio under 1 represents a decrease in the probability of NHRI adoption (a negative coefficient). The presented results show support for the hypothesis that the amount of shaming a state receives influences its decision to adopt an NHRI. The coefficient for shaming and its squared term are positive and negative, respectively, conforming to the expectation of an inverted–U relationship.\textsuperscript{28} As shaming of a state increases, it learns others’ beliefs about it’s reputation for existing in the liberal club of states, which can be a sign that its status will decrease. In order to avoid a loss of status, they will adopt an NHRI.

\textsuperscript{27} (Koo and Ramirez, 2009, 1335) also experience high collinearity in their World Polity variables causing them to include the variables separately.

\textsuperscript{28} Respective p–values: 0.072; 0.083.
The negative squared term suggests the marginal benefits of such action are highest at middling values of shaming. Those shamed very little or all of the time do not attempt to adjust their reputation using NHRI adoption.

Results associated with the controls warrant further discussion. First, both democracy and material dominance affect NHRI adoption in the manner predicted. As a state’s democracy or material dominance (holding all other variables constant) increase, the probability a state adopts an NHRI increases, to a point in which it begins to decrease again. Those states with middling scores on democracy and material dominance find marginal benefits of adoption outweigh the costs. The largest p-value associated with the variables and their squared terms is 0.017.

Interestingly, a state’s human rights practices don’t contribute to the probability that a state adopts an NHRI (p = 0.745). The result may indicate state human rights practices do not influence NHRI adoption. Perhaps states act as they will with respect to human rights, and adopt an NHRI if/when the international community calls them out using public fora.

Alternatively, the result may indicate opposite mechanisms at work with respect to human rights. Scholars propose states enter into agreements with which they already plan to comply (e.g. Downs, Rocke and Barsoom, 1996; Von Stein, 2005). Others propose gross offenders will commit and then continue to violate to signal to domestic audiences that they are in complete control (Hollyer and Rosendorff, 2011). The finding on the human rights control suggests the possibility of both of these theoretical accounts playing out around the world, thus pulling the sign of the coefficient in opposite directions resulting in no observed effects. Future research should unpack the mechanisms involved with respect to human rights practices and NHRI adoption.

How independent the judiciary is does not affect the adoption of NHRIs, either. The results suggest one of two theoretical stories. The first is that domestic judiciaries don’t have strong preferences as to whether NHRIs exist alongside them. More interestingly, however, it may be the case that some judiciaries wish for allies in legal human rights cases and some view new institutions as competition. Future work should explore the relationship between NHRIs and domestic judiciaries.

State age does not affect whether a state adopts an NHRI. Recall, I argued young states may wish to adopt NHRIs to more precisely signal to other actors their desire to be included in the global liberal regime. One may argue regime age more important than state age leveraging the
same mechanism about increasing information about a young regime. However, robustness tests show regime age has no effect on NHRI adoption, either.  

### 2.8 Conclusion

States manage their international status by taking actions to increase their reputations. They monitor what other actors say about them – international shaming – in order to estimate their reputations and then weigh the marginal benefits of adopting a norm – in this case, NHRIs – against the expected costs. The data comport the argument: increased shaming from NGOs, the media, and the UN increase the probability a state adopts an NHRI to a point, in which marginal benefits begin decreasing. Put another way, the probability a state adopts an NHRI is related to how much a state receives international shame in an inverted-U relationship.

The sudden increase in NHRI adoption throughout the world following the Paris Principles reasonably represents one of the stages of (Finnemore and Sikkink, 1998)’s norm life cycle: norm cascade. As a result of the UN’s persistence and a handful of norm entrepreneurs, the concept of an NHRI became a reality that held real potential for reputational payoffs. My results are consistent with the argument that states care about the appropriate behavior, even if in a self-interested fashion weighing the marginal gains of adoption vs. the potential costs. Preliminary analysis in other work suggests states with NHRIs receive less shaming. Research shows individuals pursue reputation and status as an end unto itself (Bakshi and Chen, 1996; Huberman, Loch and Onculer, 2004). States may act as the individuals. They may also pursue status due to increased opportunities for material benefits. The next step will explore if states receive material benefits from adoption. These two tests will further provide insight into the underlying mechanisms of the diffusion of norms along the cascade by exploring evidence for both the immediate and diffuse benefits. Either way, holding the material dominance associated with states constant, states respond to the relatively immediate, non-material reputation signals of shaming by ascribing to a norm generally accepted by the international community – adoption of an NHRI.

The present study adds to the growing literature on NHRI adoption more specifically (Cardenas, 2003; Koo and Ramirez, 2009; Pegram, 2010; Kim, 2013). (Cardenas, 2003) and (Kim, 2013)  

---

29 $p=0.47$; other results essentially unchanged except the squared term on democracy ($p=0.17$ rather than 0.02).

30 Author’s own work.
highlight how specific institutions (UN and NGOs, respectively) altered the potential reputational benefits and costs of adoption. (Pegram, 2010) offered diffusion as the mechanism for which NHRI s spread throughout the world. I add that states are immediately concerned with how other states perceive them as they attempt to position themselves in the liberal order perpetuated by the allies and the UN after WWII. Building on the argument that NHRI s diffused (Koo and Ramirez, 2009; Pegram, 2010), I add that their concern with their reputation and status drives the fact that as more states adopt, a state is more likely to adopt (Strang, 1991; Koo and Ramirez, 2009; Pegram, 2010).

As usual, the present research offers potential avenues of future study. For instance, I assume constant costs of NHRI adoption (sovereignty costs; material costs). Once states adopt an NHRI, they design the NHRI. Reading NHRI annual reports show that states opt to spend different amounts of money on establishment and operation. Other work finds the probability the NHRI can levy legal punishment (affecting sovereignty costs) decreases as the executive exerts more control over the legislature during the design process (Welch, 2015b). Although executives cannot know the exact political environment they will find themselves in the future, they must have beliefs that can formulate expected costs, thus informing how likely NHRI adoption should be. Future work should explore this dependent relationship between adoption and design.

The current study also suggests policy implications for practitioners. States acting within the norm cascade represent strategic adopters. Though some states may adopt NHRI s with sincere intentions, others may be most concerned with the immediate reputational benefits, and may seek to curb the expected sovereignty costs in the design phase. The international community helps with start-up and functioning costs and training for NHRI s. For instance, Western countries donate to Latin American and African NHRI s, Commonwealth and Scandinavian donations make up the majority of the Afghanistan Independent Human Rights Commission budget, and the UN donates to NHRI s throughout the world (e.g Nowrojee, 2001; AIHRC, 2010; Pegram, 2012; IHREC, ND). Those countries that rely on foreign funds for NHRI creation and functioning may be more subject to external monitoring and influence. If countries want to keep the funds coming in, they may make sure the adopted NHRI s meet international standards and aren’t just window dressing, minimizing adoption for short-term reputational spikes without intention of
increasing human rights respect. Future research should explore how successful donors can be in holding these countries to account.
CHAPTER 3

DELEGATING ACCOUNTABILITY: AN ANALYSIS OF NATIONAL HUMAN RIGHTS INSTITUTIONS VARIATION

The international human rights regime relies on institutions. For the past few decades scholars have focused on the effects of international (e.g. Hathaway, 2002; Hafner-Burton and Tsutsui, 2007; Hill, 2010), regional (e.g. Hawkins and Jacoby, 2010; Sikkink and Walling, 2007; Haglund, 2013), and domestic institutions (e.g. Davenport, 1995; Davenport and Armstrong, 2004; Conrad, Hill and Moore, 2014) in order to better understand the institutional functions in the regime. Scholars recognized that all three institutional contexts rely on each other to effectively constrain governments from abusing human rights with more local mechanisms making international law and norms more effective (e.g. Risse-Kappen, Ropp and Sikkink, 1999; Powell and Staton, 2009; Simmons, 2009; Lupu, 2013). Before that, practitioners realized how important domestic institutions were to the international regime as the United Nations defined and promoted National Human Rights Institutions (NHRIs): “The General Assembly [of the United Nations]…reaffirms the importance of developing...effective national institutions for the promotion and protection of human rights.”1 NHRIs are domestic institutions with a specific mandate to promote and protect human rights. And despite the international community’s decades-old call and the prevalence of NHRIs (over 100 throughout the world), scholars (especially social scientists) have been slow to take them up as objects of study (See e.g. Mertus, 2009; Goodman and Pegram, 2012, for exceptions). The few systematic social scientific studies of NHRIs focus on either their adoption (Koo and Ramirez, 2009; Kim, 2013) or their effectiveness (Cole and Ramirez, 2013). However the above quote, the first substantive point of the Paris Principles after the affirmation of the human rights regime, stresses the importance of “developing...effective national institutions (emphasis added).” For this reason I focus on the design of NHRIs. I ask, what explains NHRI design with

respect to mechanisms meant to hold violators accountable? In this way, I contribute the first quantitative study of NHRI design.

I propose a theory in which the legislature delegates a mechanism to the NHRI that holds the executive accountable – the ability to levy punishment for human rights violations. Levying punishment, such as fining and firing offenders, represents real sanctions capable of holding the executive to account (Downs, Rocke and Barsoom, 1996; O’Donnell, 1999). The executive, wishing to retain as much autonomy as possible, attempts to influence how successful the legislature delegates the ability to levy punishment to the NHRI. By taking advantage of her relative power vis-à-vis the legislature, she seeks to keep her NHRI from a design that is capable of holding her accountable for future human rights abuses. The legislature and executive, being rational actors, seek to influence the design of institutions (Koremenos, Lipson and Snidal, 2001). The legislature wishes to delegate some of its human rights tasks to the NHRI due to its specialization in human rights (Kiewiet and McCubbins, 1991; Bendor, Glazer and Hammond, 2001). The executive will attempt to influence as much of the delegation as possible by leveraging her legal and political advantages (Frye, 1997; Shugart and Mainwaring, 1997). Electoral rules, such as district magnitude and plurality elections, adjust the legislature’s willingness to thwart the executive’s attempts by incentivizing individual legislators to fight more or less with the executive over the public good of human rights provisions (Cingranelli and Filippov, 2010). The executive also takes advantage of political opportunities, such as controlling a majority or exploiting a fractionalized opposition, to exert more power in the delegation process. Ultimately, the executive may even be able to effectively hijack the whole delegation process and directly design the NHRI using unilateral measures such as decrees. Any of these scenarios that increase the executive’s relative power to the legislature during the delegation phase decrease the probability the NHRI can levy punishment. The present study contributes to the growing literature on NHRIs, as well as those on executive/legislative relationships and organizational delegation.

By estimating a Bayesian logistic regression that accounts for selection of states into the sample (whether they adopted an NHRI), I find support for four of my five hypotheses lending credence to the theoretical importance of the executive’s relative power in the delegation of accountability mechanisms to NHRIs.
3.1 A Brief History of NHRIs

NHRIs are domestic institutions created by the government of the state in which they act. They are charged with promoting and protecting human rights. They are a continuation and growth from the ombudsman model. The first ombudsman was created in Sweden in 1809. Over one hundred years later, Finland established an ombudsman (1919). The next states adopted the institution after WWII “with the expansion of state activity during and after the war [and] a new concern for protecting human rights” (Rowat, 1973, 119).²

All ombudsmen take specific complaints from the public against administrative injustice and maladministration (Rowat, 1973, 147). “The daily logs and annual reports of ombudsman offices describe a myriad of small and large individual complaints” (Zagoria, 1988, 18). Ombudsmen differ with respect to what they do with those complaints (e.g. investigate, criticize, and/or publicize). Scholars note that the “great weapon of the ombudsman is criticism” (Rowat, 1973, 49).

As ombudsmen continued to proliferate throughout the world, the atrocities of WWII increased global awareness and action against human rights abuses. Borrowing from the ombudsman model, states proposed domestic institutions to handle complaints of citizens who believed their human rights violated. States confirmed the potential importance of NHRIs in a 1946 UN Resolution and a 1978 seminar.³ However, most states did not adopt an NHRI until after 1991 when the UN International Workshop on National Institutions for the Promotion and Protection of Human Rights was held in Paris. The workshop yielded a draft specifically recommending NHRIs called the Paris Principles that were adopted by the UN General Assembly in 1993. After the Paris Principles, many states adopted an NHRI, either by establishing a new institution or by converting existing institutions such as ombudsmen (Smith, 2006).⁴

Whereas almost all NHRIs take complaints, similar to their institutional predecessors (ombudsmen), human rights complaints differ from government maladministration. Critics may contend that complaints to a government institution about governmental abuse may not lead to a change in governmental behavior even if the NHRI can investigate and/or publicly criticize

---

² Denmark(1955), Norway(1962), New Zealand(1962).
³ UN Economic and Social Resolution 9 (II), 21 June 1946; “National and Local Institutions for the Promotion and Protection of Human Rights (1978).”
⁴ Some NHRIs maintain an ombudsman structure, whereas others are organized as commissions or committees (Koo and Ramirez, 2009).
the government. Human rights abuses often happen at times critical to governmental survival or when other costs of not repressing are very high. But some NHRIs can actually levy punishment. NHRIs in states as dissimilar as Uganda and Australia have the ability to issue binding resolutions that can result in fines to the government to be awarded for victim compensation (Cardenas, 2003).

### 3.2 Explaining Variation in NHRI Accountability Mechanisms

In this section I present a theory to explain the variation in accountability mechanisms granted to NHRIs. Although past work explores adoption (e.g. Koo and Ramirez, 2009; Kim, 2013) and efficacy (Cole and Ramirez, 2013), I follow Koremenos, Lipson and Snidal (2001, 762) in recognizing that “[institutions] are the result of rational, purposive interactions among...actors to solve specific problems.” In fact, one could go as far as to say studying the effects of NHRIs without acknowledging design variation suffers from putting the cart before the horse. (Cole and Ramirez, 2013) take an important first step by looking at the effects of different types of NHRIs based on their nomenclature/leadership e.g. ombudsman vs. commission, but the more interesting variation results from NHRIs’ ability to hold violators accountable; and before exploring the effects of the design variation, we should discover what brings about those differences. I argue the variation in accountability mechanisms results from the willingness and ability of the legislature to delegate authority. Electoral rules and other actors (specifically the executive) affect legislature willingness and ability.

Legislatures are responsible for the majority of documents establishing NHRI role in government and society (Conrad, DeMeritt and Moore, 2013). The Paris Principles insist that NHRI authority “be clearly set forth in a constitutional or legislative text.” By heeding the Paris Principles, states give first mover advantage (Bendor, Glazer and Hammond, 2001) to the legislature over the executive in delegating accountability mechanisms. Therefore I start with explaining why legislatures delegate the authority to another institution. Before I start, though, I make my

---

5 In this case, the ability to levy punishment.

6 Over 75% of NHRI founding documents allocating which powers the NHRI may enact are directly from the legislature or from a constitution (Conrad, DeMeritt and Moore, 2013)

7 Although one can conceive of NHRI authority delegated by the executive in that the executive allows them to exist, this more correctly fits with the concept of adoption. Once the NHRI concept is adopted, the legislature goes to work to outline its abilities. Of course the executive will play a role in the process, which is discussed below.
assumptions explicit. The legislators making up a legislature wish to remain in power. They will attempt to do so by pursuing the policies that will increase the probability they continue their term – beholding them to citizen preferences.\textsuperscript{8} Individual citizens prefer not to be abused, and prefer others aren’t abused as well (Gronke et al., 2010; Wallace, 2013).\textsuperscript{9}

Given these assumptions, legislatures may have an interest to delegate authority to NHRIs. Since legislators wish to work towards policies that align with constituents, they should, all things being equal, wish to increase human rights protections in the state. Legislatures act as a check on the executive (Montesquieu, 1752; Malesky and Schuler, 2010; Madison, 2013) making them responsible for punishing her if she chooses to violate human rights. However, legislatures are institutions burdened with overcrowded agendas (Kiewiet and McCubbins, 1991; Bendor, Glazer and Hammond, 2001). Delegating the authority to hold executives accountable can alleviate the problem by allowing a specialized institution to shoulder some of the burden (Lupia and McCubbins, 1994). Carefully designed NHRIs represent potential restraints on violators (Cole and Ramirez, 2013). The accountability mechanism chosen matters because delegation always comes with costs. In principal–agent theory, agents possess hidden information and hidden actions (Kiewiet and McCubbins, 1991) which can lead to shirking. In this instance, “shirking” would be allowing the executive to control the NHRI making it a toothless institution. However, carefully designing the institution by delegating the ability to hold the executive legally responsible, what I call levying punishment, creates an institution that truly holds violators accountable.\textsuperscript{10}

The actor most likely affected by the accountability mechanisms delegated to the NHRI is the executive. The executive also wishes to remain in power. Although policy concessions are often the preferred method to remain in power (e.g Gandhi, 2008), repression of the opposition remains a viable tactic when the cost of concessions are too high (Duvall and Stohl, 1988; Moore, 2000; Boix, 2003; Davenport, 2007b). The executive wields the state’s monopoly on power through the military and police giving her the option to turn those weapons against the internal opposition if

\begin{itemize}
\item\textsuperscript{8} Although this assumption is more easily invoked for democracies, almost every country in the world has a legislature, even if they are the results of concessions to the opposition (Boix, 2003; Gandhi, 2008). These quasi–democratic institutions develop a surprising amount of influence (North, 1990; Boix, 2003; Malesky and Schuler, 2010).
\item\textsuperscript{9} Although see (Conrad, Hill and Moore, 2014) for an argument that the majority isn’t concerned with minority groups being subjected to torture.
\item\textsuperscript{10} I find (O’Donnell, 1999)’s definition of horizontal accountability most useful for conceptualizing accountability: “The existence of state agencies that are legally enabled and empowered, and factually willing and able, to take actions that span from routine oversight to criminal sanctions or impeachment in relation to actions or omissions by other agents or agencies of the state that may be qualified as unlawful.”
\end{itemize}
deemed necessary (Moore, 2010). In this way, the executive is ultimately responsible for human rights violations within her borders.\footnote{(DeMeritt, 2015) relaxes the assumption of the executive as a unitary actor allowing for delegation issues within the executive branch. I opt to conceptualize the executive as a unitary actor for parsimony reasons. Even if delegation within the branch is important, the head executive remains ultimately responsible.} For these reasons, the executive will work to influence the delegation process.

In order to explain the relationship between the executive and legislature during the delegation of accountability mechanisms to the NHRI, I follow an electoral bargaining approach (Frye, 1997). According to the model, the more powerful the executive in relation to the legislature, the easier will she affect the design of institutions.\footnote{(Frye, 1997) also includes uncertainty in his electoral bargaining model. I find uncertainty distracting from the main point of executive–legislative relationships I wish to explore, but do control for uncertainty as described below. Future studies should explore the effect of uncertainty at a more theoretically rich level with respect to NHRI design.} Executive power flows from two sources: constitutional provisions and party control (Shugart and Mainwaring, 1997). Executives cannot act unilaterally (Bueno de Mesquita et al., 2003; Howell, 2003); they must best navigate the legal and political environment in which they find themselves.

Constitutional laws, specifically electoral laws, determine who comprises the legislature. The executive often relies on her power to persuade (Neustadt, 1960), which can be easier or harder depending on the legislative make–up. Electoral rules affect legislator incentives to pursue policies, such as human rights protections (Cingranelli and Filippov, 2010), by determining how responsive legislators have to be to constituent preferences (Powell, 2000; Persson and Tabellini, 2003). One such rule is district magnitude. District magnitude refers to the number of representatives elected in a set geographical space. As the number of elected officials in the space increase, district magnitude increases. As district magnitude increases, individual legislator action becomes harder to monitor by constituents in turn creating a collective action problem among the legislators. Any individual can claim responsibility for public goods, so no one individual has an incentive to produce those goods. Good human rights represent a public good, so we can expect that as legislators become less incentivized to produce public goods, they will fight less to pursue policies such as delegating true accountability mechanisms to NHRI s. As legislators try less, the executive’s relative power grows and she influence delegation with higher probability resulting in an NHRI without the ability to levy punishment so that she may retain the ability to repress if deemed necessary.
District Magnitude Hypothesis: As district magnitude increases, the probability the NHRI can levy punishment decreases.

Besides how many legislators govern a district (district magnitude), the determination of winners at election time also influences the relative power of the executive and legislature. Plurality rules, when the winner is determined by who receives the most votes and that winner takes all, are associated with fewer public goods and more targeted redistribution if the public good is not particularly desirable (Lizzeri and Persico, 2001). Since plurality rules force a winner–take–all system, individual legislators offer public goods with the largest appeal, selecting a few important issues and setting aside less important ones. This begs the question – do constituents find human rights particularly desirable? Do they warrant legislator effort? (Cingranelli and Filippov, 2010, 246) assert constituencies deem human rights less important than issues such as “terrorism or immigration.” However, (Wallace, 2013) surveys U.S. citizens asking about their views on torture. Even when terrorism is an imminent threat, the majority of respondents prefer the government not to torture suggesting human rights issues may be at least on par with terrorism. Also survey data from Europe suggests human rights are one of the most important issues to citizens (e.g Commission of the European Communities, 1989). Taking these data to suggest citizens consider human rights important, legislators have incentive to fight harder for human rights provisions, reducing the probability the executive influences the delegated design of the NHRI. Therefore, we should expect that plurality rules will allow the legislature to delegate accountability mechanisms such as granting the ability to levy punishment to the NHRI.

Plurality Hypothesis: Countries with plurality rules have a higher probability of containing an NHRI that can levy punishment compared to countries without plurality rules.

Executive relative power also derives from party control. The executive enjoys more strategic positioning in the legislature when she can appeal to her party support. (Kingdon, 1989) presents an illustrative U.S. example in which Congressmen were asked who played a role in their decision making during the Nixon administration. Forty–two percent of Republicans spontaneously answered the president, while only 12 – 14% of democrats answered the same way. The example highlights the importance of the executive in using her power in the legislative body. The executive can appeal to members of her own party using various strategies that increase the benefits

13 (Edwards, 1989) notes these appeals are even more effective with party leaders.
of legislator compliance (or increase the cost of non–compliance). Politicians share both pasts and futures together. The executive can appeal to the past and future by calling up favors, promising favors, or threatening sanctions (Edwards, 1989). These strategies work best on members of her own party as they share political pasts and futures.

How do these executive strategies affect delegation to NHRI s to levy punishment? The executive wishes to keep as much autonomy as possible to make the decisions necessary to stay in power, including repressing. Since the executive’s strategies to affect her power in the legislature succeed with a higher probability with her own party, we have insight into when she will be most successful. When her party has a majority, the probability these party–control strategies – personal appeals, offering favors, threatening sanctions – increase her chances to influence the delegation of accountability mechanisms to the NHRI.

Majority Hypothesis: When the executive’s party controls a majority of the legislature, the probability an NHRI can levy punishment decreases.

The majority hypothesis suggests the executive most successfully influences delegation when she has a majority because she can lobby allies. Implicit to this argument is there exists a set of legislators not sharing a party with the executive – the opposition. The opposition often prefers to work against the executive for the same reasons partisans wish to work for her – executive outcomes affect their political futures. When the executive cannot affect successes, the opposition succeeds. However, often multiple parties represent the opposition. These parties exhibit different levels of agreement with each other. A fractionalized opposition allows the executive advantages to influencing legislative policy. The less cohesive the opposition, the easier the executive can sway members of the opposition parties to her side increasing her support, thus increasing the probability she influences policy. Since the executive wishes to retain autonomy to repress if necessary, she will look to take advantages of fissures in the opposition when the time to delegate authority to the NHRI comes.

Fractionalization Hypothesis: As the opposition becomes more fractionalized, the probability an NHRI can levy punishment decreases.

Thus far the theoretical discussion presented assumes the legislature delegates authority to the NHRI. Although the Paris Principles state NHRI s should be established by legislative or constitutional means, the executive may be able to hijack the delegation process. In fact, just under
20% of NHRI founding documents originate from the executive branch in the form of executive decrees (Conrad, DeMeritt and Moore, 2013) (see Figure 3.1). If able to do so, the executive attains the highest relative power vis-à-vis the legislature, as she is now the delegator. Executives monopolizing the delegation process amounts to capturing the first-mover advantage in Principal–Agent theory (Bendor, Glazer and Hammond, 2001) wherein the actor that acts first increases the probability of attaining preferences. Wrestling the delegation process comes with costs (in the form of legitimacy) given the international community’s stated preference of legislature and constitutional delegation, but rational actors pursue policies when perceived benefits outweigh perceived costs. If the executive believes she can act unilaterally without domestic recourse (Waterman, 2009) (costs of acting unilaterally low) or the need for repression is high enough (costs for not acting unilaterally high), she will do so with higher probability.\textsuperscript{14} Bypassing bargaining with elements of the legislature and keeping the delegation process to herself allows the executive to design the NHRI closest to her preferences, with weaker accountability mechanisms.

\textsuperscript{14}Future research should address when this is the case.
Decree Hypothesis: If an NHRI is founded by executive decree, the probability the NHRI can levy punishment decreases.

3.3 Research Design

The proposed hypotheses are conditional upon a country having adopted an NHRI, which has repercussions for the modeling strategy. In order to account for the selection into the sample, I first estimate a standard probit including predictors found to be important in state decision to adopt NHRI: global dominance and its squared term, level of democracy and its squared term, level of shaming and its squared term, state age, human rights respect, and judicial independence (Welch, 2015a). I then take the propensity score for whether a state will have an NHRI in a given year and use it as a control variable in the second stage.

In the second stage, I estimate a Bayesian\textsuperscript{15} logistic regression on a cross-section of the global population of NHRI according to the International Coordinating Committee of National Institutions for the Promotion and Protection of Human Rights (ICC)\textsuperscript{16} from 1991–2012.\textsuperscript{17} I model the choice to delegate the ability to levy punishment to an NHRI. Therefore I estimate the binary choice model:

\[
y_i \sim Bern(p_i), \quad p_i = \lambda(X_i\beta + Z_i\gamma)
\]

Where \( \lambda \) is the inverse logit function, \( i \) indexes NHRI, \( y_i \) is the dichotomous dependent variable whether the NHRI has authority to hold the executive accountable,\textsuperscript{18} \( X_i \) is a matrix of covariates, and \( \beta \) is a vector of parameters to be estimated; \( Z_i \) is a matrix of latent variable covariates and \( \gamma \) is a vector of the associated parameters to be estimated. For each iteration of the

\textsuperscript{15}Maximum likelihood estimators are normally distributed about the true mean with the smallest possible variance in large samples. However, these properties do not always hold in smaller samples, such as mine (Casella and Berger, 2002; Train, 2009). Instead of relying on asymptotic standard errors and the assumption of normality, I use MCMC simulation to sample from the posterior distribution of the coefficients.

\textsuperscript{16}The ICC is a group of NHRI that coordinates the relationship between NHRI and the UN and accredits NHRI based on UN standards. For more information and a list of participating NHRI see http://nhri.ohchr.org/EN/Contact/NHRIs/Pages/Global.aspx.

\textsuperscript{17}I choose 1991 as the beginning year because of the importance of the Paris Principles in defining NHRI as a separate institution.

\textsuperscript{18}Operationalization is discussed below.
simulation, $Z_i$ is drawn from its estimated posterior distribution.\textsuperscript{19} I fit the model with R2JAGS, generate random initial values for two simulated chains of length 500000 with the first 10000 iterations discarded. I confirmed convergence visually,\textsuperscript{20} and numerically.\textsuperscript{21} Given the evidence across a large number of convergence diagnostics, I accept the model as converging.

### 3.3.1 First Stage Variables

**Dependent Variable.** The dependent variable is dichotomous taking the value 1 if an NHRI is legally adopted in a particular country in a particular year, and 0 otherwise. I coded the adoption year by consulting various sources – NHRI annual reports, NHRI and government websites, third–party websites, and academic pieces. Figure 3.2 presents the global distribution of NHRIs in 2012. Those shaded grey or black have an NHRI. Those shaded black have an NHRI with the legal ability to levy punishment.

**Independent Variables.** I include the variables found to be important in (Welch, 2015a): global dominance and its squared term, level of democracy and its squared term, level of shaming and its squared term, state age, human rights respect, and judicial independence. I measure global dominance and level of shaming (and, thus their squared terms) with a latent variable model.\textsuperscript{22} For level of democracy I include a latent variable created by (Pemstein, Meserve and Melton, 2010). I create a state age variable by subtracting the year a state was founded from the current year in the database. I obtain the founded years from (Gleditsch and Ward, 2013). I control for human rights respect using a latent variable measure from (Fariss, 2014) constructed using a dynamic modeling process with many of the most used measures used to measure human rights (e.g. PTS (Gibney and Dalton, 1996), CIRI (Cingranelli and Richards, 2012), ITT (Conrad, Haglund and Moore, 2013)). Lastly, I control for the independence of the judiciary including a latent variable created by (Linzer and Staton, 2011) that leverages eight other judicial measures from past scholarship.

\textsuperscript{19}Specifically, instead of a set data matrix $X$, every iteration uses (potentially) different values for $Z$ drawn from a normal distribution for each data point centered at the estimated mean with an estimated standard deviation: $Z_i \sim \mathcal{N}(\mu_{\text{estimated}}, \sigma_{\text{estimated}})$. This strategy incorporates the uncertainty in the latent variables. The results are not substantively different for the coefficients of interest than if the mean of the latent variable estimate was used as the set values. The model with included uncertainty also has a lower DIC (82.3 vs 84.6). The results for the model without uncertainty are included in the Appendix.

\textsuperscript{20}The trace plots, autocorrelation plots, and Geweke–Brooks plots are included in Appendix.

\textsuperscript{21}The majority of R-hats are 1.001 or 1.002, with one at 1.003 and 1.004 respectively.

\textsuperscript{22}For in–depth descriptions of these latent variables see (Welch, 2015a).
Figure 3.2: NHRIs that can Levy Punishment, 2012. White=No NHRI, Gray=NHRI, no Punishment, Black=NHRI, Punishment
3.3.2 Second Stage Variables

Dependent Variable. My theory is that the the greater the executive’s relative power to the legislature, the lower the probability that the NHRI will contain provisions that allow holding the executive to account. “Accountability implies not only answerability, but also the legal obligation to answer or the institutionalized right of an agent of accountability to impose sanctions on public officials” (Mainwaring, 2003, 7, emphasis in the original). NHRI design varies leaving different institutions with different abilities including publishing reports, investigating, and levying punishment. Whereas almost every NHRI is legally able to publish reports or investigate, far fewer have the ability to levy punishment (see Figure 3.3). Using the above definition as a guide I focus on the NHRI ability to levy punishment because it includes active sanctioning (Downs, Rocke and Barsoom, 1996; Dunn, 1999; Keohane, 2002).\textsuperscript{23} Effective accountability requires sanctioning such as those activities allowed to the NHRI s that levy punishments such as fining and firing violators. The low variation on the low-sanctioning activities and the theoretical conceptualization in which “accountability itself requires penalties beyond publicity” (Kenney, 2003, 64) steer my decision to focus on levying punishment as the best representation of an accountability mechanism.

I operationalize the dependent variable as a dichotomous indicator of whether the NHRI design includes the ability to levy punishment. I obtained the data from the National Human Rights Institution Data Collection Project Conrad et al. (2013). Data were collected using content analysis on various documents including published reports by NHRI s, reports by actors monitoring NHRI s, government and NGO websites, and academic pieces.

Independent Variables. To evaluate the District Magnitude, Plurality, Majority, and Fractionalization hypotheses, I use data from the Database of Political Institutions (Beck et al., 2001). For the District Magnitude Hypothesis, I use the mean district magnitude of the House and Senate, calculated as the “weighted average of the number of representatives elected by each constituency size, if available” (Keefer, 2012, 16). If not available, the number of seats are divided by the number of constituencies. For the Plurality Hypothesis I use a dichotomous indicator for whether legislators win election using winner-take-all/first-past-the-post rules. To evaluate the Majority Hypothesis I

\textsuperscript{23} Also, exploratory factor analysis separates levy punishment on another dimension than publishing or visiting sites (what I view as high-sanctioning vs. low-sanctioning dimensions).
The Number of NHRIs Able to Perform Each Activity

Figure 3.3: Permitted Activities of NHRI. Source: (Conrad, DeMeritt and Moore, 2013)
use a dichotomous indicator of whether the executive’s party holds a majority in the lawmaking houses of a state. For the Fractionalization Hypothesis I use an index of the fractionalization of the opposition party in the legislature. The index is a continuous variable from 0 to 1 calculated as the “probability that two deputies picked at random from among the opposition parties will be of different parties” (Keefer, 2012, 11).

To evaluate the Decree Hypothesis, I use a dichotomous indicator of whether the NHRI’s founding documents originated from the direct executive authority such as executive decrees. The variable takes a value of 1 if so, and a 0 otherwise. The measure is taken from the National Human Rights Institutions Data Collection Project (Conrad et al., 2013).

<table>
<thead>
<tr>
<th>Table 3.1: Design Independent Variables</th>
</tr>
</thead>
<tbody>
<tr>
<td>Variable</td>
</tr>
<tr>
<td>---------------------------</td>
</tr>
<tr>
<td>District Magnitude</td>
</tr>
<tr>
<td>Plurality</td>
</tr>
<tr>
<td>Majority</td>
</tr>
<tr>
<td>Fractionalization</td>
</tr>
<tr>
<td>Decree</td>
</tr>
</tbody>
</table>

3.3.3 Control Variables

I control for the level of judicial independence due to its position in the “web of accountability” in a state (Mainwaring, 2003, 15). Due to the legal sanctioning activities of the judiciary, it may be either complementary or competitive with NHRIs. I remain agnostic as to which relationship is more common, as accounts of both exist (e.g Carver, 2012; Pegram, 2012; Peruzzotti, 2012; Welch, 2015a). Depending on the potential relationship with the judiciary, one may expect the probability of levying punishment either increasing or decreasing. The presence of a strong judiciary also influences direct executive action as more independent judiciaries act as stronger checks on the executive (e.g Powell and Staton, 2009; Whittington, 2009). I operationalize judicial independence with a latent measure by (Linzer and Staton, 2011) taking continuous values from zero to one (with one most independent). Latent variables allow the researcher to incorporate the uncertainty of the measurement estimates of unobservable (or latent) qualities – in this case
judicial independence. I include the uncertainty for this latent measure and every other that I am able to use in the model by allowing the model to select from an estimated posterior (normal) distribution of the latent variable for each draw, rather than a point estimate (e.g. mean).

The ideology of the executive may also influence her decision to stymie the design of NHRI with respect to punishment power. More liberal executives can reasonably be believed to allow the probability of levying punishment to increase, whether to satisfy their electoral base or in the expectation that they won’t violate human rights as often (also beneficial for the electoral base). The more liberal an executive, the higher the probability the NHRI can levy punishment. I operationalize the extent of the executive’s liberalness with an ordinal ideology orientation variable from the Database of Political Institutions (Beck et al., 2001) ranging from 0 to 3, with 3 being the most liberal.

I control for level of democracy because democracies exhibit more liberal ideals (e.g Howard and Donnelly, 1986), willingness to compromise (e.g Gurr, 1986a; Henderson, 1991), respect for rule of law (e.g Valentino, Huth and Balch-Lindsay, 2004), and ultimately respect human rights more often (e.g. Poe and Tate, 1994; Davenport, 1995). I operationalize democracy using the Unified Democracy Scores (UDS) – scores estimated from a Bayesian statistical measurement model from 10 existing measures (Pemstein, Meserve and Melton, 2010). UDS scores are continuous and range from -2 to 2.

I include the level of human rights in the given state. Abuse of human rights strongly predicts future abuse of human rights (e.g Conrad and Moore, 2010). If the executive expects to need to utilize repression in the near future, her valuation of an NHRI without the ability to hold her accountable increases. I expect worse human rights practices to be correlated with lower probability of ability to levy punishment. I operationalize human rights respect with a latent measure of human rights from (Fariss, 2014) that, much like the judicial independence and democracy scores used, uses a Bayesian measurement model incorporating different human rights measures and dynamic effects of time. The latent human rights score is continuous and ranges from around -2 to 4.

I control for the uncertainty in human rights levels. As uncertainty increases, politicians may design institutions that more strongly constrain decision-making (Frye, 1997). For example in post-communist Eastern Europe, the uncertainty of the political environment led to current lead-
ers pursuing policies and institutions that constrained the executive as a way to “lock in” future politicians if they sought to stray from the liberal path of human rights protection (Moravcsik, 2000). I measure the uncertainty about human rights practices as the standard deviation of the estimate for the human rights score from (Fariss, 2014). The standard deviation captures the uncertainty in measurement and the uncertainty of the time dynamics and thus acts as a proxy for human rights uncertainty.24

Lastly, I control for the estimated probability that a state has an NHRI, which I obtain from the saved propensity scores from the first stage equation. This controls for the selection associated with the dataset including only states that have NHRIs (see Stone, 2008, for a similar strategy with IMF conditionality given the state applied for an IMF grant.).

Table 3.2: Design Control Variables

<table>
<thead>
<tr>
<th>Variable</th>
<th>Central Tendency</th>
<th>Std. Dev.</th>
<th>n</th>
</tr>
</thead>
<tbody>
<tr>
<td>Judicial Independence</td>
<td>0.49</td>
<td>0.27</td>
<td>194</td>
</tr>
<tr>
<td>Executive Liberalness</td>
<td>1.31</td>
<td>1.30</td>
<td>201</td>
</tr>
<tr>
<td>Democracy</td>
<td>0.33</td>
<td>0.79</td>
<td>212</td>
</tr>
<tr>
<td>Human Rights</td>
<td>0.28</td>
<td>1.28</td>
<td>201</td>
</tr>
<tr>
<td>Human Rights Uncertainty</td>
<td>0.25</td>
<td>0.07</td>
<td>201</td>
</tr>
<tr>
<td>Pr(NHRI)</td>
<td>0.52</td>
<td>0.18</td>
<td>191</td>
</tr>
</tbody>
</table>

3.4 Results

Table 3.3 presents parameter estimates (means), standard deviations, 95% credible intervals,25 and the effective n.26 The results mostly support my theory that the executive’s ability to influence the delegation of accountability mechanisms to NHRIs affects overall design. When constitutional provisions create large district magnitudes, the probability an NHRI can levy punishment decreases suggesting legislators lack incentive to fight for public goods such as human rights when constituents cannot easily assign responsibility for policy actions.

24 Better measures should be developed. However, given the lack of one, I deem the proxy appropriate as a control measure.
25 Bayesian credible intervals can be interpreted exactly as frequentist confidence intervals for the sake of hypothesis testing due to my choice of uninformative priors.
26 Effective n is a measure of effective sample size after simulation.
Table 3.3: Delegation of NHRI Accountability Mechanism, 1991-2012

<table>
<thead>
<tr>
<th>Hypotheses</th>
<th>Mean</th>
<th>Std. Dev.</th>
<th>95% Cred. Interval</th>
<th>Effective N</th>
</tr>
</thead>
<tbody>
<tr>
<td>District Magnitude</td>
<td>-0.17</td>
<td>0.06</td>
<td>(-0.30, -0.06)</td>
<td>2000</td>
</tr>
<tr>
<td>Plurality</td>
<td>0.07</td>
<td>0.90</td>
<td>(-1.64, 1.74)</td>
<td>1300</td>
</tr>
<tr>
<td>Majority</td>
<td>-2.21</td>
<td>0.97</td>
<td>(-4.23, -0.41)</td>
<td>1000</td>
</tr>
<tr>
<td>Fractionalization</td>
<td>-6.55</td>
<td>2.22</td>
<td>(-11.31, -2.67)</td>
<td>2000</td>
</tr>
<tr>
<td>Decree</td>
<td>-3.27</td>
<td>1.64</td>
<td>(-6.74, -0.36)</td>
<td>2000</td>
</tr>
<tr>
<td>Judicial Independence</td>
<td>-1.27</td>
<td>0.52</td>
<td>(-2.37, -0.35)</td>
<td>630</td>
</tr>
<tr>
<td>Executive Liberal</td>
<td>0.72</td>
<td>0.39</td>
<td>(-0.02, 1.51)</td>
<td>2000</td>
</tr>
<tr>
<td>Democracy</td>
<td>-1.53</td>
<td>0.84</td>
<td>(-3.21, 0.03)</td>
<td>2000</td>
</tr>
<tr>
<td>Human Rights</td>
<td>-5.64</td>
<td>2.39</td>
<td>(-11.08, -1.78)</td>
<td>2000</td>
</tr>
<tr>
<td>Uncertainty</td>
<td>11.51</td>
<td>6.78</td>
<td>(-1.40, 25.71)</td>
<td>900</td>
</tr>
<tr>
<td>Pr(NHRI)</td>
<td>3.52</td>
<td>3.08</td>
<td>(-2.06, 9.65)</td>
<td>470</td>
</tr>
</tbody>
</table>

However, when legislators obtain seats with plurality voting rules, the probability the NHRI can levy punishment is unaffected. The argument I make hinges on two moving parts. First, plurality rules must restrict legislators to pursue only a few policy areas. Given the restriction, they should focus on those policy areas most important to the largest number of people. I drew the conclusion that human rights rank in the top of importance based on survey data from the U.S. and Europe (Commission of the European Communities, 1989; Wallace, 2013), but use a global sample. Perhaps, people in countries such as the U.S. and those of Europe possess postmaterial preferences making human rights concerns relatively important, as opposed to those that may focus on security, food, or the economy (Inglehart, 1997). Future studies should look to tease apart the importance of human rights in public opinion and how that affects legislator behavior.

When the executive exerts party control either by leaning on her majority in the legislature or taking advantage of a non–cohesive opposition, her relative power increases allowing her to affect the design of the NHRI. Both coefficients were negative as expected with the 95% of the posterior distribution also negative (Majority $\beta$, standard deviation: $-2.15$, 0.95; Fractionalization $\beta$, standard deviation: $-5.42$, 2.07).
When the executive wrests delegation away from the legislature, she successfully designs NHRIs in accordance to her preferences. I estimate the coefficient for the Decree Hypothesis at -3.93 with standard deviation of 1.55. The results for this hypothesis along with the others (except for the Plurality Hypothesis) lend support to the theory proposed that the executive seeks to influence the delegation process to NHRIs as much as possible in order to design an institution allowing her the most autonomy in decision making.

The coefficients for the controls warrant brief discussion. As a state’s judiciary increases in independence, the probability an NHRI can levy punishment decreases. Courts are the classic institution for protecting individual rights. Perhaps states with well-functioning judiciaries do not see a reason to delegate the ability to levy punishment to another institution. Future research should further explore the relationship between the courts and NHRIs.

As the executive becomes more liberal, the probability an NHRI can levy punishment increases (almost 95% of the parameter distribution is positive). Similarly, as uncertainty increases about human rights practices, those designing the institutions seek to limit the uncertainty by delegating more accountability authority. Interestingly, as human rights respect in a state increase, the probability an NHRI can levy punishment decreases. Perhaps states with good human rights practices design NHRIS with different goals in mind, such as turning their attention outward. For instance, in December 2004, The German Institute for Human Rights held a workshop for a number of NHRIs from Europe and Central Asia (German Institute for Human Rights, 2004), an example of an NHRI looking to solve international human rights problems rather than domestic human rights problems. The fact that different NHRIs are designed and consequently focus on different issues presents an interesting avenue of future research. The coefficient estimate for democracy contained zero in the 95% credible interval, suggesting more democratic states are no more or less likely to delegate the ability to levy punishment.

Even though the coefficient on the propensity score used to control for selection (Pr(NHRI)) is not significant at standard levels, I prefer the selection–corrected model to that without the propensity control for theoretical reasons and based on the deviance information criterion (DIC). The DIC is a measure of model fit that penalizes extra parameters, much like the AIC or BIC, used to evaluate Bayesian model fit (Spiegelhalter et al., 2002). The DIC for the selection–controlled
model is 85.1 compared to the model without the propensity score included, 130.0 (lower DIC is preferred).

The estimated parameters from the regression give an idea of which variables predict designing an NHRI with the power to levy punishment, and with what certainty. However, due to the non-linearity in outcome probabilities and the fact that some of the explanatory variables are binary, I present first differences to better understand the substantive meaning of the estimates (Long, 1997; King, Tomz and Wittenberg, 2000; Brambor, Clark and Golder, 2006).27

The first differences28 shed light on which variables are most important (see Table 3.4).29 As expected from the 95% credible interval, plurality rules don’t affect the probability of levying punishment. As district magnitude goes from its lowest to highest values, the probability the NHRI can levy punishment drops by 34%. When the executive has a majority in both houses of the legislature the probability the NHRI can levy punishment drops by 26%. Substantively, the opposition going from complete solidarity to completely at odds results in a decrease in probability that the NHRI can levy punishment by 34%. Lastly, despite the relatively large coefficient for the Decree coefficient (−3.95), a change from no executive decree to executive decree as the source of the NHRI’s allowed activities decreases the probability of levying punishment by only 13%, which while lower than the other variables is still a substantive effect.

Table 3.4: Design First Differences

<table>
<thead>
<tr>
<th>Variable</th>
<th>Mean</th>
<th>Std. Dev.</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>District Magnitude</td>
<td>-0.34</td>
<td>0.12</td>
<td>(-0.59, -0.14)</td>
</tr>
<tr>
<td>Plurality</td>
<td>-0.09</td>
<td>0.11</td>
<td>(-0.34, 0.11)</td>
</tr>
<tr>
<td>Majority</td>
<td>-0.26</td>
<td>0.13</td>
<td>(-0.53, -0.02)</td>
</tr>
<tr>
<td>Fractionalization</td>
<td>-0.34</td>
<td>0.18</td>
<td>(-0.70, -0.03)</td>
</tr>
<tr>
<td>Decree</td>
<td>-0.13</td>
<td>0.08</td>
<td>(-0.33, -0.02)</td>
</tr>
</tbody>
</table>

27 I held all continuous values at their means and categorical values at their medians. Variable of interest changes from minimum value in data to maximum value.

28 I used Zelig to generate first differences from a Bayesian logit (King, Tomz and Wittenberg, 2000; Goodrich and Lu, 2007) with the same specifications as above except Zelig uses one chain rather than two and uses the Metropolis–Hastings algorithm instead of the Gibbs sampler. Results are substantively similar to the main model and are presented in Appendix.

29 It is important to remember that the current project analyzes cross-sectional data, thus the first differences show the result of a theoretical change in a variable of interest for an average state.
3.5 Conclusion

The evidence presented supports the claim that executives use legal and political advantages to adjust their relative power in order to influence NHRI design. The more power the executive can enact, the lower the probability the NHRI will have the ability to hold the executive to account via levying punishment.

Rational actors know institutions can affect their decision calculus, thus they seek to design institutions in line with their preferences (Koremenos, Lipson and Snidal, 2001). Scholars find design affects a wide range of outcomes including trade (e.g Milgrom, North and Weingast, 1990; Kono, 2009), alliances (e.g Poast, 2012; Benson, Bentley and Ray, 2013), and duration of peace (e.g. Long, Nordstrom and Baek, 2007; Badran, 2014), to name a few. This paper contributes to the small literature on institutional design with respect to human rights institutions and the even smaller literature on NHRI design, by exploring what factors influence different design decisions. It also contributes to the little systematic social science work on NHRIs (Koo and Ramirez, 2009; Cole and Ramirez, 2013; Kim, 2013).

As is often the case, the current project suggests a number of future avenues of study. A logical next step is to explore how NHRI design variation effects outcomes. For instance, the actors in my theory act on the belief that NHRIs that levy punishment more effectively hold executives accountable. Future research should test the assumption.

In past work, I’ve found that NHRIs keep executives from torturing if they have ratified the UN Convention Against Torture (Welch, 2015c). The theoretical mechanism is one of information gathering and dissemination due to the secretive nature of torture. Future research should explore the variation in other design dimensions, such as information accumulation and sharing.

Whereas the current research elucidates what predicts delegation of accountability mechanisms, broadly, future research should theorize about states with different characteristics. For example, if the legislature/executive relationship predicts NHRI ability to punish, variation in legislative traditions should be further explored. What are the differences between presidential and parliamentarian systems? Healthy vs. tainted legislative environments e.g. democracies, hybrid regimes, and autocracies?

The present research and proposed research on NHRIs is of utmost importance to the international human rights regime’s priority stated in the Paris Principles preamble that “priority
should be accorded to the development of appropriate arrangements at the national level to ensure effective implementation of international human rights standards.”
APPENDIX A

SUPPLEMENTARY MATERIAL FOR CHAPTER 2

A.1 Latent Variable Models

As is common with many of the concepts in political science, many of the variables for this study are constructs that cannot be observed directly. Instead of directly measuring them, I use available data that I assume are manifestations of the latent quantity. One can think of it as having observable data $y$, and using them to discover some latent quantity $x$ (Treier and Jackman, 2008, 201). This approach allows me to recover each country’s level of the latent concept (e.g. shaming, dominance) as an unknown parameter allowing for distributions as well as point estimates. The model is as follows:

Let $i = 1, \ldots, n$ country-years, $j = 1, \ldots, m$ latent concepts, and $k = 1, \ldots, K_j$ observed indicators.

$$y_{i,j,k} \sim N(\mu_{i,j,k}, \tau_{j,k})$$

$$\mu_{i,j,k} = \beta_{j,k} X_{i,j}$$

$$\beta_{j,k} \sim N(0, 10)$$

$$\tau_{j,k} \sim Gamma(1, 1)$$

I estimated the models in a Bayesian framework with Markov Chain Monte Carlo simulations (10000 iterations and 1000 burn–in) using the R2Jags package in R to call JAGS. I used uninformative priors for the dominance model. For the shaming model I used a more informative prior that assumed any shaming from any of the actors (media, NGOs, UN) contributed a positive amount of shaming to the state. The choice of prior for the shaming variable is necessary for convergence and theoretically justifiable, as it does not make sense that states experience less shame from shaming episodes. Visual inspection of trace plots (Gill, 2002; Lam, ND), autocorrelation plots
(Lam, ND), Gelman–Rubin–Brooks Plots (Brooks and Gelman, 1998), and Geweke–Brooks plots (Geweke, 1992) suggest convergence.

A.2 State Dominance Figures

Fifty Most Dominant States

Figure A.1: Fifty Most Materially Dominant States
Figure A.2: Fifty Least Materially Dominant States
APPENDIX B

SUPPLEMENTARY MATERIAL FOR CHAPTER 3

B.1 Convergence Statistics

I present the visual convergence diagnostics for the reader’s discretion.

B.1.1 Trace Plots

Trace plots are time-series of the chains. Each parameter shows good mixing. If the parameters did not mix well, the chains would get stuck in localized areas of the parameter space or show upward/downward trends (Gill, 2002; Lam, ND).

B.1.2 Autocorrelation Plots

Convergence exhibits itself as the values of autocorrelation decreasing over draws (Lam, ND). The plots show all parameters quickly approaching around zero.

B.1.3 Geweke–Brooks Plots

Geweke–Brooks plots display the Z-score of the Geweke–Brooks diagnostic. The Geweke–Brooks diagnostic test convergence by assessing a test for equality of means of the first and last part of the Markov chains. Convergence occurs when the null hypothesis that samples are drawn from the stationary distribution of the chain is confirmed (Geweke, 1992). Therefore, convergence should show z scores between -2 and 2.
Figure B.1: Trace Plots
Figure B.2: Trace Plots II
Figure B.3: Autocorrelation Plots
Figure B.4: Autocorrelation Plots II
Figure B.5: Autocorrelation Plots III
Figure B.6: Geweke–Brooks Plots
Figure B.7: Geweke–Brooks Plots II
Figure B.8: Geweke–Brooks Plots III
### B.2 Other Models

#### Table B.1: Design Model with No Uncertainty

<table>
<thead>
<tr>
<th>Hypotheses</th>
<th>Mean</th>
<th>Std. Dev.</th>
<th>95% Cred. Interval</th>
<th>Effective N</th>
</tr>
</thead>
<tbody>
<tr>
<td>District Magnitude</td>
<td>-0.13</td>
<td>0.06</td>
<td>(-0.25, -0.03)</td>
<td>1100</td>
</tr>
<tr>
<td>Plurality</td>
<td>-0.91</td>
<td>0.86</td>
<td>(-2.69, 0.74)</td>
<td>2000</td>
</tr>
<tr>
<td>Majority</td>
<td>-2.15</td>
<td>0.95</td>
<td>(-4.04, -0.37)</td>
<td>730</td>
</tr>
<tr>
<td>Fractionalization</td>
<td>-5.42</td>
<td>2.07</td>
<td>(-9.80, -1.56)</td>
<td>2000</td>
</tr>
<tr>
<td>Decree</td>
<td>-3.93</td>
<td>1.55</td>
<td>(-7.49, -1.25)</td>
<td>2000</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Controls</th>
<th>Mean</th>
<th>Std. Dev.</th>
<th>95% Cred. Interval</th>
<th>Effective N</th>
</tr>
</thead>
<tbody>
<tr>
<td>Judicial Independence</td>
<td>-2.36</td>
<td>4.00</td>
<td>(-10.31, 5.09)</td>
<td>2000</td>
</tr>
<tr>
<td>Executive Liberal</td>
<td>0.62</td>
<td>0.33</td>
<td>(-0.01, 1.27)</td>
<td>2000</td>
</tr>
<tr>
<td>Democracy</td>
<td>-1.09</td>
<td>1.44</td>
<td>(-3.93, 1.71)</td>
<td>2000</td>
</tr>
<tr>
<td>Human Rights</td>
<td>-1.59</td>
<td>0.64</td>
<td>(-2.97, -0.46)</td>
<td>2000</td>
</tr>
<tr>
<td>Uncertainty</td>
<td>24.12</td>
<td>7.122</td>
<td>(11.25, 39.50)</td>
<td>2000</td>
</tr>
<tr>
<td>Pr(NHRI)</td>
<td>2.53</td>
<td>2.94</td>
<td>(-3.19, 8.57)</td>
<td>1400</td>
</tr>
</tbody>
</table>

#### Table B.2: Design Zelig Results

<table>
<thead>
<tr>
<th>Hypotheses</th>
<th>Mean</th>
<th>Std. Dev.</th>
<th>95% Cred. Interval</th>
</tr>
</thead>
<tbody>
<tr>
<td>District Magnitude</td>
<td>-0.10</td>
<td>0.05</td>
<td>(-0.20, -0.02)</td>
</tr>
<tr>
<td>Plurality</td>
<td>-0.61</td>
<td>0.71</td>
<td>(-2.00, 0.79)</td>
</tr>
<tr>
<td>Majority</td>
<td>-1.51</td>
<td>0.72</td>
<td>(-2.96, -0.13)</td>
</tr>
<tr>
<td>Fractionalization</td>
<td>-2.88</td>
<td>1.38</td>
<td>(-5.71, -0.26)</td>
</tr>
<tr>
<td>Decree</td>
<td>-2.73</td>
<td>1.22</td>
<td>(-5.37, -0.53)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Controls</th>
<th>Mean</th>
<th>Std. Dev.</th>
<th>95% Cred. Interval</th>
</tr>
</thead>
<tbody>
<tr>
<td>Judicial Independence</td>
<td>-1.59</td>
<td>2.16</td>
<td>(-5.74, 2.56)</td>
</tr>
<tr>
<td>Executive Liberal</td>
<td>0.40</td>
<td>0.28</td>
<td>(-0.12, 0.96)</td>
</tr>
<tr>
<td>Democracy</td>
<td>-0.44</td>
<td>0.93</td>
<td>(-2.30, 1.36)</td>
</tr>
<tr>
<td>Human Rights</td>
<td>-0.92</td>
<td>0.46</td>
<td>(-1.86, -0.07)</td>
</tr>
<tr>
<td>Uncertainty</td>
<td>4.97</td>
<td>2.95</td>
<td>(-1.01, 10.65)</td>
</tr>
<tr>
<td>Pr(NHRI)</td>
<td>1.86</td>
<td>1.84</td>
<td>(-1.64, 5.52)</td>
</tr>
</tbody>
</table>
REFERENCES


78


IHREC. ND. “NHRI Capacity Development Partnership Project.”. URL: http://www.ihrec.ie/international/nhricapacitydevelopment.html


URL: https://fsu.academia.edu/RyanWelch?c.p=t

URL: https://fsu.academia.edu/RyanWelch?c.p=t


URL: http://data.worldbank.org/indicator

BIOGRAPHICAL SKETCH

Ryan Merrill Welch was born June 23, 1982 to Dwight and Amy Welch. He received an IB diploma from Stanton College Preparatory High School. He attended University of Florida (Go Gators!) where he received a B.A. in Political Science; and both a B.S. and M.S. in Entomology. When he decided to go back to school, he traded his house flies for human rights. He will receive his Ph.D. in Political Science from Florida State University.