Estimating the Effects of Human Rights Treaties on State Behavior

Daniel W. Hill Jr. Florida State University

Though research suggests that international regimes that coordinate economic and security policy can alter state behavior, research examining the effect of human rights treaties on state behavior has found that these agreements do little to curb repressive practices. However, those studies neglect to account for the fact that several of the state-level characteristics which are known to affect repressive practices also influence the likelihood of a state making a formal commitment to the human rights regime. States that commit and states that do not are likely to have different domestic institutional features. Systematic heterogeneity across ratifiers and nonratifiers makes it difficult to infer the level of repression that would have been observed in a state had it not committed to the treaty in question. This paper employs matching techniques that address this problem and allow for more valid inferences about the effects of human rights treaties on repressive practices. The impact of human rights treaties is examined in the context of three of the five core UN human rights treaties. The results are quite interesting: ratification of the CAT and, to a lesser extent, the ICCPR is associated with reduced respect for physical integrity rights while ratification of the CEDAW has a positive impact on observance of women's rights. These findings suggest that more treaty-specific theory building is needed.

1The data and other files needed to replicate the analysis below will be made available at http://myweb.fsu.edu/dwh06c/ on publication. An online appendix with supplementary materials may be found at www.cambridge.org/cjo/XXXX.

2Leeds (2003) examines the the effect of military alliances on states’ decisions to aid their allies during wartime; Simmons (2000) and Simmons and Hopkins (2005) examine the effect of the adoption of IMF article VIII on international monetary policy.

© Southern Political Science Association, 2010
ISSN 0022-3816
ratify a human rights treaty in the first place (Hathaway 2007) adds further concern. This means that many statistical models of human rights violations that include treaty dummies contain covariates that are not only correlated, but causally related to one another. Those states that have ratified and those that have not are likely to differ systematically from one another in a number of ways, and many of these differences are likely to affect human rights practices. Unless we directly confront the process by which states select themselves into these treaties, we may be drawing biased inferences about the effects they have on observance of basic rights.³

It is difficult to construct a convincing counterfactual when the typical state that ratifies a treaty is far more likely to do so than the typical nonratifier. It is also difficult to determine what the effect of these treaties is when the typical ratifier possesses more of the state-level characteristics known to be associated with good human rights practices than the typical nonratifier. Ideally analysis to determine what, if any, effect human rights treaties have on actual practices would be conducted on a set of states that are nearly identical to one another in every way save some have ratified the treaty and some have not. One solution to the inability to randomly assign a treaty “treatment” to countries is to use matching techniques to approximate such a counterfactual comparison. This study does so. Several studies examining the effect of participation in international institutions have recognized the need to account for self-selection processes (e.g., Abouharb and Cingranelli 2006, Landman 2005, Przeworski and Vreeland 2000, von Stein 2005), and one study has identified matching as an especially appropriate solution to this challenge (Simmons and Hopkins 2005). However, the utility of these techniques is not widely recognized, and this study echoes Simmons and Hopkins (2005) while presenting new results pertaining to the human rights regime and its effects on state behavior.

In the next section I review existing explanations of why human rights treaties should, or should not, be expected to affect state practices, and the empirical work that has accompanied such explanations. I next present my research design, which employs matching techniques to produce a more appropriate sample with which to estimate the effects of human rights treaties on state behavior. Matching procedures allow me to estimate quantities that should be of special interest to scholars studying human rights treaties, in particular the sample average treatment effect among the treated (SATT; Ho et al. 2007, Morgan and Winship 2007), i.e., the effect that a human rights treaty has had among the states that have ratified it and for which there exists an appropriate group of nonratifying states to compare them to. To determine whether such an effect exists I examine human rights behavior in the context of three of the core UN human rights treaties: the International Covenant on Civil and Political Rights (ICCPR), the Convention against Torture and other Cruel, Inhuman or Degrading Treatment or Punishment (CAT), and the Convention on the Elimination of All Forms of Discrimination against Women (CEDAW). Analysis is conducted using pooled time-series data on 165 states from 1976 to 2006.

International Enforcement

From the standpoint of scholars studying international institutions, why a human rights treaty would have any effect on human rights practices is not clear at first glance. Compared with other international institutions, the formal enforcement mechanisms in the international human rights regime are rather weak. Committees and councils created to determine whether states are complying with a particular treaty frequently operate on a system of self-monitoring whereby states submit reports on human rights violations taking place within their borders and the measures they are taking to eliminate them (Hafner-Burton 2005; Lebovic and Voeten 2006). Aside from the obvious lack of incentive to report one’s own wrongdoings, even if states felt compelled to call attention to their own transgressions the human rights committees that view the reports are powerless to punish recalcitrant states. Many treaties provide no recourse against offending states beyond “naming and shaming” by other state parties, and these complaints are often not made public.⁴

³To be clear, I am not talking about sample selection bias. Measures of repression are readily available for nearly all states, so observability of the outcome is not an issue. The problem is that treaty ratification is non-random; certain types of states are more or less likely to select into human rights treaties than others and this self-selection process must be accounted for in statistical analysis.

⁴In fact, though article 22 of the CAT and the second optional protocol to the ICCPR allow for their respective committees to hear complaints from individuals, a state may ratify the treaties and declare itself not bound by these provisions. This means that many ratifying states only need to worry about the possibility of other states filing complaints against them, not their own citizens. The CEDAW allows for complaints to be filed by state parties only, so there is no possibility that the committee would receive information about violations from citizens.
Some scholars may not worry that formal enforcement mechanisms within the human rights regime are weak; after all, the majority of the work on state behavior vis-à-vis international institutions has focused primarily on how informal enforcement mechanisms can alter the behavior of recalcitrant states. Unfortunately, the usual set of tenable self-enforcement mechanisms are simply not suitable in the context of the human rights regime (See Simmons (2009) for a thorough discussion). The shadow of the future and fear of reciprocal violation cannot induce compliance in the area of human rights, since it is not clear that states have anything to gain by jointly observing human rights or anything to lose should each (or any) party fail to do so. States should hardly be expected to behave strategically in the realm of human rights, which is to say that their human rights behavior is not the result of expectations that other states are, or are not, going to commit human rights violations (Koremenos 2007). Incentives to violate human rights do not arise from the nature of interactions in the international arena but rather from circumstances at the domestic level.

Another informal enforcement mechanism thought to operate in international regimes is fear of damage to one’s reputation (Keohane 1984; Lipson 1991; Simmons 2000). In theory, once states have made formal, public commitments to obey the rules of the regime noncompliance may result in a loss of credibility that is costly enough to deter violations. This mechanism is similar to fear of reciprocal violation in that it depends on the violating state expecting to be deprived of something in the future, namely any number of international agreements it could make with other states had it only proven itself to be trustworthy. In practice, however, it is unlikely that states will be shunned by potential partners because of a bad human rights record. Noncompliance in one area of international law does not necessarily signal an inability or unwillingness to comply in other areas, and this may be especially true of noncompliance with human rights regimes (Downes and Jones 2002).

Some scholars argue that better enforcement is not the proper remedy for widespread noncompliance. In this view, violations are usually not intentional so punishing them will not create compliance. Rather, violations are the result of ambiguous directives, inability (rather than unwillingness) to comply, and a lag between treaty adoption and implementation (Chayes and Chayes 1993). This argument appears untenable in the context of the human rights regime. While provisions contained in human rights treaties are often general and broad in scope, it seems implausible that states could unwittingly violate the human rights of their citizens. The basic tenets and even specific provisions of the human rights regimes are relatively simple to understand. Arguing that ambiguity leads to violations would be more plausible in the context of regimes with esoteric rules and regulations, such as international trade or monetary regimes. Similarly, it is difficult to argue that most states lack the ability to stop repressive practices. Expecting states to tighten fiscal policy upon receipt of an IMF loan may be unreasonable in some cases; expecting them to stop imprisoning and torturing political opponents does not seem too excessive. Temporal delays also do not appear to explain noncompliance with human rights regimes. Hafner-Burton and Tsutsui (2007) find that among states exhibiting low levels of compliance at the time of ratification the passage of time does not improve practices. In short, enforcement is a glaring problem for the human rights regime.

**Normative Pressure**

If it is unreasonable to expect that the weak enforcement mechanisms created by the human rights regime will induce compliance, perhaps there is some hope in the socialization process described by constructivist scholars. It has been argued that when a sufficient number of states have adopted an international norm a kind of peer pressure emerges that can coerce other states into adopting the norm, resulting in a “norm cascade” (ostensibly the pressure to conform builds with each state that adopts the norm, the end result being that the norm becomes universal or nearly so; Finnemore and Sikkink 1998; Hawkins 2004). The concept of legitimacy is intimately connected to the norm cascade process. States are compelled to accept norms in order to show that they are not deviants, i.e., they recognize themselves as part of an international community of states and deserve the recognition and privileges that their peers receive. Failure to abide by norms concerning human rights can result in condemnation from human rights INGO’s that is troublesome enough to change state practices in some instances (Keck and Sikkink 1998). Indeed, Hafner-Burton and Tsutsui (2005) and Landman (2005) find that greater NGO membership

---

5Though there may be cases where states are unable to reign in the behavior of non-state actors (e.g., militias), it seems unlikely that the state would be unable to stop the behavior of government agents. It also seems unlikely that problems of this sort explain the majority cases of non-compliance with the human rights regime.
among the citizens of a state decreases the likelihood of repression.

Concerns about international legitimacy are also thought to influence the adoption of norms through their relation with legitimacy at the domestic level (Finnemore and Sikkink 1998). It is argued that states that fail to abide by prevalent international norms risk losing the support of their citizenry, since citizens in modern states are likely to evaluate the performance of their government relative to the performance of other governments. Disturbed by the fact that their state is failing to provide a level of rights protection commensurate with that of other states, citizens will be less likely to comply with their government’s directives hence the government’s position in power will be weakened.

The primary appeal of this argument stems from the fact that, intuitively, it seems that observance of basic human rights would be one of the most important factors in determining a government’s legitimacy. If citizens of a nation value anything they are likely to value the right to live free from gross discrimination, physical harm, and fear of death at the hands of the state. Surely this would be high on the list of criteria people use to evaluate the performance of their government relative to others. However, if we look closer at the relationship between legitimacy and repression it becomes less clear that concerns about losing support will prompt states to adopt better practices. In one analysis, instances of large-scale repression (particularly violent repression) are indicative of a breakdown of the state’s authority (Arendt 1970). This would seem to imply that states resorting to repressive measures have already lost some degree of legitimacy (i.e., citizen support) and are using force to maintain stability. Since leaders’ concerns about legitimacy are ultimately concerns about their position in power, adopting the preceding viewpoint would lead one to believe that repressive leaders have simply substituted force for lost legitimacy and are resigned to rule without the support of a large portion of their citizenry. Leaders unsure of their hold on power are likely to repress their populations, and if doing so leads to further decreases in societal support and more acts of dissent repression becomes even more likely (Young 2008).

Furthermore, some leaders’ tenures may not depend much on societal support to begin with. Bueno de Mesquita et al. (2003) argue that the amount of support leaders require to stay in power may be quite small under certain institutional arrangements. We would not expect these leaders to change their behavior because of concerns about legitimacy since increases in societal support are likely to be orthogonal to the likelihood of remaining in office. Repressive leaders lacking citizen support should not be expected to try and regain support by curbing repressive measures since they are likely to perceive the use of force as an expedient means of maintaining control (Gurr 1988).

Should international peer pressure to adopt global human rights norms be increased by the ratification of a human rights treaty? Ratification would certainly seem to indicate that a state has at least acknowledged the value of the principles contained in the treaty. If it does not signal that the state has fully adopted norms concerning human rights, it should certainly give extra leverage to those imploring the state to do so (i.e., human rights activists, NGOs, etc.). If there is any reason to expect that human rights treaties will curb state repression, perhaps it is because of the additional pressure from NGOs that states invite when they ratify. Still, some feel that treaties will not contribute to this pressure but rather alleviate it, offering an opportunity for states to increase their repressive behavior. In this account, the presence of dense networks of NGOs within a state will increase the amount of information about violations available to both citizens and non-state actors and make compliance more likely regardless of whether a state has ratified any human rights treaties, while ratification could potentially result in higher levels of repression (Hafner-Burton and Tsutsui 2005).

**Domestic Enforcement**

Besides providing leverage for INGOs who wish to make state violations public, some have argued that treaty ratification may strengthen domestic factors that constrain state behavior. Helfer and Slaughter (1997) expect that treaty ratification in democratic states will empower domestic groups who will pressure their leaders to fulfill their international legal obligations, human rights treaties being no exception. Similarly, Hathaway (2002) expects (and finds) that democracies are more likely than other states to comply with the human rights treaties they have ratified, a phenomenon she attributes to the mobilization of the public mentioned by Helfer and Slaughter, but also to the fact that democracies are likely to have strong legal institutions which provide citizens with legal recourse in the event that their rights are violated.6 While it is questionable whether

---

6Neumayer (2005) finds the same pattern.
or not democratic publics will actually mobilize to constrain their governments during periods of repression (Davenport, Moore, and Armstrong 2008), judiciaries can constrain state behavior where the public values state compliance with the judiciary, that is where the public has become willing to withdraw support from a noncompliant government (Vanberg 2005).

Although judiciaries can restrict state behavior, we should not be too quick to put our faith in courts as a solution to compliance problems in the human rights regime. Powell and Staton (2009), like Hathaway, consider courts an effective enforcement mechanism for international legal obligations. However, they find that states with strong judiciaries are likely to avoid commitment to human rights treaties and violate their provisions. This raises an important point: domestic enforcement mechanisms are likely to influence whether or not a state ratifies a treaty in the first place. For example, Goodcliffe and Hawkins (2006) find a link between the rule of law and commitment to the CAT, while Hathaway (2007) finds that the more democratic a state is, the more likely the state is to ratify several human rights treaties, including the three examined here. Another important point is that the institutions that are viewed as viable enforcement mechanisms are also likely to be related to a state’s propensity to repress its citizenry. Since institutional arrangements impact a state’s decision to ratify a human rights treaty, the states that have ratified a given human rights treaty are likely to have domestic features that differ sharply from states that have not ratified that treaty. The fact that these same features also affect human rights practices poses a problem when trying to determine the effect of human rights treaties on state behavior (von Stein 2005). If domestic-level features determine who ratifies, and if this selection process creates systematic differences between ratifiers and nonratifiers then our ability to infer what a ratifying state’s level of repression would have been had it not ratified the treaty is considerably compromised (if one uses standard estimation techniques such as those used in previous analyses to determine treaty effects) (Dehejia and Wahba 2002).

This is not merely a point about empirical analysis, it is a theoretical one that has been made by several authors with respect to international institutions in general and the human rights regime in particular (Powell and Staton 2009; Simmons and Hopkins 2005; von Stein 2005). As discussed above, theories of commitment to the human rights regime and theories of repression adopt many of the same explanatory factors. It would certainly be difficult to make the case that a state’s choice to ratify a human rights treaty is independent of its choice to violate human rights (Powell and Staton 2009). Since many of the same institutional features are front and center in theories of ratification and theories of state behavior, it is important to explicitly address this connection in our tests of such theories. I am not offering a new theory explaining how human rights treaties affect state behavior. Rather, I am arguing that existing theories of compliance with the human rights regime imply that empirical tests which do not confront the process by which states commit to the regime in the first place and expressly consider how this process affects compliance are flawed. The arguments discussed above have conflicting expectations, and the question of whether human rights treaties affect better practices is still open for empirical investigation. If we are to accurately determine the effect of a human rights treaty on a ratifying state’s practices, we must compare these states with nonratifying states that are nearly identical to the ratifying states on all dimensions, including institutional characteristics, that affect states’ decisions to engage in repressive behavior and to commit to the human rights regime. Building on work by Simmons and Hopkins (2005), the analysis below employs matching techniques which make such a comparison possible and thus allows for more valid inferences about the effects of human rights treaties on state behavior.

**Research Design**

The data employed in the analysis are time-series cross-sectional data on 165 states, covering the years 1976–2006. The unit of observation is the country-year. The temporal domain coincides with the entry into force of the ICCPR, which was the first human rights treaty to enter into force of the three examined here. Analysis for the other two treaties begins the year they enter into force, which is 1980 for the CEDAW and 1986 for the CAT.

**Estimation Technique/Dependent Variables**

Since the three treaties being examined here are intended to protect different kinds of rights it would not be appropriate to use the same dependent variable when determining their effects. Therefore, a different model is estimated for each treaty. The type
of model estimated varies according the nature of the variable used to measure states’ respect for the type of right in question.

**ICCPR**

Though the ICCPR is designed to protect a broad range of human rights (including the right to self determination, a fair trial, free movement across borders, freedom of religion and free speech), I follow other scholars in choosing to focus on the most egregious kind of violations, namely violations against personal integrity rights (Poe and Tate 1994; Poe, Tate, and Camp-Keith 1999; Camp-Keith 1999). The reason for this decision is that it would be impossible to develop an indicator to measure all of the rights covered by the ICCPR. Articles 6, 7, and 9 of the ICCPR prohibit extralegal killings, torture, and arbitrary arrest and detention, respectively. Respect for these rights will be measured using the Cingranelli-Richards (CIRI) physical integrity rights scale.\(^7\) The scale ranges from 0 to 9, with higher values indicating greater respect for physical integrity rights. Given that this variable has a fairly large range, I estimate parameters using OLS with robust standard errors clustered by country.

**CAT**

The CAT, as its name implies, is designed to prohibit states from using torture, torture being defined as the infliction of severe physical or mental pain or suffering for interrogation or punishment purposes. In order to capture the amount of torture occurring in a state, data were again taken from the CIRI human rights database. The database contains an ordinal variable which measures the state’s respect for the right to freedom from torture. The values of the variable range from 0 to 2, where a value of 0 indicates torture occurred “frequently,” a value of 1 indicates that torture occurred “occasionally,” and a value of 2 indicates that no instances of torture occurred. Since the torture measure is an ordinal scale, I estimate an ordered probit model, again with robust standard errors clustered by country.

**CEDAW**

Past research on human rights violations have used measures of respect for women’s rights that are either very narrow or too broad and inclusive.\(^8\) As an indicator for observance of women’s rights, I draw more variables from the CIRI human rights database. I use three ordinal scales from the CIRI data, one measuring women’s social rights, one measuring women’s economic rights and another measuring women’s political rights. All three variables are ordinal scales that range from 0 to 3. Values of 0 indicate that women lack legal provision of the rights in question, while values of 3 indicate that women are afforded these rights in both law and practice. Middling values indicate that women are granted rights legally but these rights are restricted in practice. Each scale was created with specific provisions of the CEDAW in mind.\(^9\) Three separate models will be estimated to determine what effect ratification of the CEDAW has on women’s rights. I use an ordered probit model to estimate the effects of the CEDAW on respect for each category of women’s rights, again clustering standard errors by country.

**Covariates of Repression**

Besides a treaty variable (coded 1 for every country year in which the treaty being examined is (or has been) ratified and 0 otherwise) each model will contain control variables to account for other factors that have been shown to affect states’ repressive practices. Over the years scholars have managed to build a list of “usual suspect” predictors of state repression. I control for these and other factors in each of the three analyses.\(^10\) The measures for each concept are discussed separately below.

Many studies have found a positive relationship between democracy and the protection of human rights (Davenport 1995, 1999, 2007; Henderson 1991; Poe and Tate 1994; Poe, Tate, and Camp-Keith 1999). There are several posited causal mechanisms by which democracy leads to observance of human rights, most of them centering on constraints on leadership; democratic leaders who are accountable to the public and other government actors will find it more difficult than their autocratic counterparts to

---

\(^7\)For a detailed description of this measure and its properties see the online description at http://ciri.binghamton.edu/.

\(^8\)For example, Hathaway (2002) proxies women’s political rights using the proportion of men serving in the legislature and Hafner-Burton and Tsutsui (2005) use the Political Terror Scale (Gibney, Cornett, and Wood 2009).

\(^9\)In particular, the political rights scale pertains to articles 1, 3, 7, and 8; the economic rights scale pertains to articles 10 and 11; the social rights scale pertains to articles 9, 10, 12, 13, 14, 15, and 16.

\(^10\)Each model contains the same covariates save the treaty dummies.
EFFECTS OF HUMAN RIGHTS TREATIES

The presence of external (international) conflict is also the onset or continuation of a civil war and otherwise. Country-years where the state in question experienced civil war involvement is taken from the PRIO/UPPSALA armed conflict data (Gleditsch et al. 2002). The civil war variable is coded 1 during threats with repression (Davenport 1995, 2007). A number of INGOs operating in the state (Hafner-Burton and Tsutsui 2005; Neumayer 2005). Greater numbers of INGOs in a state are thought to be indicative of stronger links to global civil society, which improves human rights practices by increasing the amount of information available to domestic and external actors, thereby making governments more vulnerable to public criticism from human rights activists (Hafner-Burton and Tsutsui 2005). There is also some evidence that higher numbers of INGOs make a state more likely to ratify a human rights treaty; Hathaway (2007) finds that higher numbers of INGOs make a state more likely to ratify the CAT. Data on NGO membership is taken from Hafner-Burton and Tsutsui (2005). Their measure is a count of how many INGOs citizens of the state claim membership with.

State involvement in internal conflicts seems to affect repressive behavior (Poe and Tate 1994; Poe, Tate, and Camp-Keith 1999). Empirical work has shown that governments nearly uniformly respond to threats with repression (Davenport 1995, 2007). A variable indicating civil war involvement is taken from the PRIO/UPPSALA armed conflict data (Gleditsch et al. 2002). The civil war variable is coded 1 during country-years where the state in question experienced the onset or continuation of a civil war and 0 otherwise. The presence of external (international) conflict is also controlled for using PRIO/UPPSALA armed conflict data, as the presence of such conflicts has also been linked to the use of repression (Poe and Tate 1994).

Another state-level characteristic that has recently been found to affect a state’s use of repression is the number of INGOs in the state (Hafner-Burton and Tsutsui 2005; Neumayer 2005). Greater numbers of INGOs in a state are thought to be indicative of stronger links to global civil society, which improves human rights practices by increasing the amount of information available to domestic and external actors, thereby making governments more vulnerable to public criticism from human rights activists (Hafner-Burton and Tsutsui 2005). There is also some evidence that higher numbers of INGOs make a state more likely to ratify a human rights treaty; Hathaway (2007) finds that higher numbers of INGOs make a state more likely to ratify the CAT. Data on NGO membership is taken from Hafner-Burton and Tsutsui (2005). Their measure is a count of how many INGOs citizens of the state claim membership with.

As a measure of judicial independence, I use the International Country Risk Guide’s (ICRG) rule of law index (International Country Risk Guide 2005). The rule of law index ranges from 0 to 6, with higher scores indicating “a strong court system,” and low scores indicating “a tradition of depending on physical force or illegal means to settle claims” (International Country Risk Guide 2005). While the rule of law index does have problems, namely that it includes components other than strength of state’s court system, it is currently the only available measure of judicial independence.

11See Cross 1999, Camp-Keith (2002b) and Powell and Staton (2009) for some exceptions.
Matching

I now describe the intuition behind the matching process that allows me to draw more valid inferences about the effects of human rights treaties on respect for human rights. Scholars who wish to know the effect of international institutions on state behavior must pose a counterfactual, that is they must ask whether the state’s behavior would have been the same had it not joined the institution in question. Ultimately we are interested in comparing some outcome of interest across two different conditions. In this case, we are interested in the difference between a ratifying state’s level of repression and the level of repression that would be observed had that state not ratified the treaty. That is, we would like to know: \( \delta_i = y_i^{\text{Treaty}} - y_i^{\text{NoTreaty}} \) where \( y_i^{\text{Treaty}} \) is state \( i \)’s level of repression under the condition where it has ratified the treaty, \( y_i^{\text{NoTreaty}} \) is state \( i \)’s level of repression under the condition where it has not ratified the treaty, and \( \delta_i \) is the difference in state \( i \)’s level of repression across these two conditions. The problem, of course, is that at any particular moment in time we will not be able to observe a state that has ratified the treaty in the counterfactual condition where it has not ratified the treaty. Nor will we be able to observe a state that has not ratified the treaty in the counterfactual condition where it has ratified the treaty. This impossibility has been termed the fundamental problem of causal inference (Holland 1986).

One way to estimate the effect of a treaty would be to compare levels of repression in ratifying states with levels of repression in nonratifying states. The problem with this, as mentioned previously, is that domestic-level characteristics are almost certain to influence a state’s decision to join a human rights treaty, that is ratification is nonrandom. Democracy, NGO membership and the strength of a state’s judiciary are all likely to affect a state’s decision to ratify. This means that ratifying states are likely, on average, to have very different characteristics from non-ratifying states. These characteristics are also likely to affect the state’s use of repression against its citizenry. Some scholars have already pointed out that this problem of self-selection in the realm of international legal obligations could bias our estimates of the effects of international institutions on state behavior. von Stein (2005) advocates the use of Heckman selection models (Heckman 1979) to correct for this potential bias, but the approach used here seems more appropriate in this case. The problem we face with respect to international institutions is one of self-selection. States that commit are different from states that do not, and these differences contribute to the decision to select into a “treatment.” These differences also contribute to their decision to repress their citizens, which makes it difficult to separate the effect of the treaty from the effect of the institutional features that led them to ratify. When we can observe many of the characteristics that lead states to ratify (e.g., democracy, NGOs, judicial strength) and incorporate these into our analysis via matching we can sidestep the econometric problems associated with selection models (Simmons and Hopkins 2005). Since theories of commitment to the human rights regime posit selection into treaties based on many of the observable characteristics one would normally include in a statistical model of state repression, matching seems a better choice than a selection model.

Given the presence of this problem, if we are to determine the effect of ratification on the use of repression among states that have ratified some treaty, we must compare them to states that are nearly identical in every way save they have not ratified the treaty. Ideally we would limit our analysis to include states that have identical values on all covariates save the treaty in question. In practice, however, we can only approximate this ideal situation since no two states have identical values on all of the covariates included in the analysis. One relatively simple way to ensure that the states being compared to each other have similar values on all covariates except the treatment variable is to recode the data so that covariates which may assume many values (e.g., continuous variables) are grouped into several similar categories, perform exact matching using the recoded data, and then transform the variables back to their original values. For example, IR scholars often treat states which receive Polity democracy scores higher than some arbitrary threshold as “consolidated” or “strong” democracies. Rather than performing exact matching on the Polity score (i.e., matching only observations which have identical values), one could recode Polity so that states with a value of 7 or higher (strong democracies) are considered appropriate matches for each other (provided, of course, that they are similar on all other covariates as well), as are states with values between 1 and 6 (weak democracies), between −6 and 0 (weak autocracies), and between −10 and −7 (strong autocracies). This procedure is called coarsened exact matching because variables used in the matching procedure are “coarsened” (i.e., recoded so that they assume fewer values) prior to the matching procedure, which matches exactly on the recoded variables (Iacus, King, and Porro 2008).
To perform CEM one must decide which values of the variables used in the analysis should be grouped together. Recoding for CEM was done as follows: the Polity score was recoded exactly as described above, the ICRG index was grouped into 3 values (one for observations receiving scores from 0 to 2, another for scores of 3 to 4, and another for scores of 5 to 6), and the INGO variable was grouped into bins ranging from 0 to 10, 11 to 50, 51 to 100, 101 to 500, 501 to 1000, and 1001 and above. The natural logs of GDP per capita and population were recoded according to the quintiles of their empirical distributions (i.e., observations in the first 20% of the distribution of each variable were grouped together, as were observations falling in the second 20%, and so on). Variables measuring the presence or absence of internal and external conflict are dichotomous and so could not be “coarsened” any further. Table 1 shows how the matching procedure resulted (generally) in better balance between the treatment and control groups for each treaty. The standardized differences between treatment and control group means in the raw and matched datasets are shown for each covariate. Also shown are the variance ratios for each covariate.\footnote{Differences and variance ratios are shown for the imputation in which the difference in means between the treatment and control groups was largest after matching and for which the variance ratio was farthest from 1.}

Matching procedures allow me to accomplish a counterfactual comparison by “trimming down” the sample of states so that the “control” (i.e., non-ratifiers) and “treatment” (i.e., ratifiers) groups are balanced on all other covariates in the model (e.g., the distribution of the Polity democracy score among ratifiers in the matched samples should be very similar to that of nonratifiers). In this way we can come as close as possible to estimating the counterfactual level of repression for a state that entered the treaty, that is the expected value/probability of repression under the counterfactual condition of nonratification. Estimates of the sample average treatment effect for the treated may then be estimated using procedures described below. This quantity is of significant interest in the context of human rights treaties. Substantively, the SATT is the average change in human rights behavior among the group of states that have ratified each treaty (and for which we can find a suitable group of nonratifying states to compare them to) that is associated with the treaty in question.

As with any cross-national dataset, there were many observations with missing values. The missing data problem was particularly acute in the dataset constructed here; most of the variables used in the models below had at least several hundred missing values, and in some cases the number of missing values was more than well over 1,000 (out of 4,946 observations).\footnote{The proportion of missing values for each variable can be found in the online appendix at http://journals.cambridge.org/JOP/DOI.} Rather than use listwise deletion (i.e., dropping those observations with missing values) missing values were imputed using the Amelia II program (Honaker, King, and Blackwell 2009). Unless there are relatively few observations with missing values and it can be demonstrated that the observations with missing values are a random subset of the sample, listwise deletion can bias inferences (Little and Rubin 1987). No formal tests were performed to determine whether data were missing completely at random; cross-national datasets are notorious for missing data for nonrandom subsets of states, so it is highly unlikely that data were not missing systematically.

In order to avoid matching observations using imputed values, imputation was performed after observations were matched according to the observed values of the covariates. Observations were first grouped together according to their pattern of missingness, and within these groups CEM was conducted using the observed values. Five imputed datasets were created from the final matched dataset and each model below was estimated using each of the five imputed datasets. Results from the analyses were pooled using methods presented in Rubin (1987).

The sample average treatment effect for the treated (SATT) was calculated for each treaty. Coefficients for the other covariates are not shown here, as we are primarily interested in the effect of the treaties on state behavior.\footnote{Tables containing coefficient estimates for the other covariates can be found in the online appendix.} The other covariates are included in the models merely to control for any remaining imbalance that may exist between treatment and control groups after matching. Were we to properly estimate their effects we would need to redefine the treatment and perform matching again. Treaty coefficients are reported
Table 1 Balance after Matching, Standardized Differences in Means and Variance Ratios

<table>
<thead>
<tr>
<th>Covariate</th>
<th>CAT Diff. before</th>
<th>CAT Diff. after</th>
<th>ICCPR Diff. before</th>
<th>ICCPR Diff. after</th>
<th>CEDAW Diff. before</th>
<th>CEDAW Diff. after</th>
</tr>
</thead>
<tbody>
<tr>
<td>External War</td>
<td>.204</td>
<td>Dropped</td>
<td>.102</td>
<td>.025</td>
<td>.083</td>
<td>.026</td>
</tr>
<tr>
<td>Civil War</td>
<td>.069</td>
<td>.014</td>
<td>.081</td>
<td>.038</td>
<td>.061</td>
<td>.012</td>
</tr>
<tr>
<td>Democracy</td>
<td>.828</td>
<td>.084</td>
<td>.903</td>
<td>.233</td>
<td>.649</td>
<td>.288</td>
</tr>
<tr>
<td>INGOs</td>
<td>.91</td>
<td>.393</td>
<td>.712</td>
<td>.378</td>
<td>.648</td>
<td>.462</td>
</tr>
<tr>
<td>Judiciary</td>
<td>.561</td>
<td>.082</td>
<td>.288</td>
<td>.208</td>
<td>.139</td>
<td>.272</td>
</tr>
<tr>
<td>GDP per capita (logged)</td>
<td>.39</td>
<td>.088</td>
<td>.307</td>
<td>.091</td>
<td>.017</td>
<td>.255</td>
</tr>
<tr>
<td>Population (logged)</td>
<td>.358</td>
<td>.102</td>
<td>.274</td>
<td>.067</td>
<td>.291</td>
<td>.221</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Covariate</th>
<th>Ratio before</th>
<th>Ratio after</th>
<th>Ratio before</th>
<th>Ratio after</th>
<th>Ratio before</th>
<th>Ratio after</th>
</tr>
</thead>
<tbody>
<tr>
<td>External War</td>
<td>.532</td>
<td>Dropped</td>
<td>.532</td>
<td>.778</td>
<td>.597</td>
<td>.761</td>
</tr>
<tr>
<td>Civil War</td>
<td>.887</td>
<td>.96</td>
<td>.873</td>
<td>.903</td>
<td>.902</td>
<td>.962</td>
</tr>
<tr>
<td>Democracy</td>
<td>.691</td>
<td>.889</td>
<td>1.07</td>
<td>.957</td>
<td>.837</td>
<td>.923</td>
</tr>
<tr>
<td>INGOs</td>
<td>3.35</td>
<td>1.7</td>
<td>3.42</td>
<td>1.99</td>
<td>2.54</td>
<td>1.62</td>
</tr>
<tr>
<td>Judiciary</td>
<td>1.01</td>
<td>.918</td>
<td>1.16</td>
<td>1.07</td>
<td>1.02</td>
<td>.934</td>
</tr>
<tr>
<td>GDP per capita (logged)</td>
<td>1.07</td>
<td>.92</td>
<td>.973</td>
<td>.98</td>
<td>1.05</td>
<td>1.06</td>
</tr>
<tr>
<td>Population (logged)</td>
<td>.706</td>
<td>.805</td>
<td>.631</td>
<td>.772</td>
<td>.856</td>
<td>.871</td>
</tr>
</tbody>
</table>

(Table 2) for models using the raw data and for models using the matched, imputed data. The main quantities of interest here are not coefficients for the treatment variables (the treaties), rather they are predicted values (or probabilities) of the dependent variable for the treatment group, counterfactual predicted values for this group (for the treatment group this would be the expected value of repression if the treaty were not ratified), and the (average) difference between these values. The results from the three models are quite interesting; effects vary quite a bit across treaties, ranging from beneficial to deleterious.

**ICCPR**

While the results of past research examining the effects of the ICCPR are mixed, the results here support those who have found it to have a slightly negative effect on state practices (Hafner-Burton and Tsutsui 2005). The results from this analysis are found in Table 2. Though the SATT was calculated for this treaty it is not reported here to conserve space, and also because it would not be very interesting as it is very close to the coefficient reported in the second row of Table 2. While this effect is negative (indicating a decrease in observance of physical integrity rights) and significant, it is substantively rather small before matching and becomes even more so after the imbalance between treatment and control groups on the other covariates in the model is corrected. These results certainly indicate that states are not complying with their obligations under the ICCPR, however they do not indicate that states are egregiously violating the

Table 2 Treaty Coefficients

<table>
<thead>
<tr>
<th>Treaty</th>
<th>Dep. Var.</th>
<th>Matching</th>
<th>Coefficient</th>
<th>s.e.</th>
<th>p</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>ICCPR</td>
<td>Phys. Int.</td>
<td>None</td>
<td>-.429</td>
<td>.191</td>
<td>.027</td>
<td>1650</td>
</tr>
<tr>
<td>ICCPR</td>
<td>Phys. Int.</td>
<td>CEM</td>
<td>-.405</td>
<td>.101</td>
<td>.001</td>
<td>1643</td>
</tr>
<tr>
<td>CAT</td>
<td>Torture</td>
<td>None</td>
<td>-.538</td>
<td>.09</td>
<td>.000</td>
<td>1658</td>
</tr>
<tr>
<td>CAT</td>
<td>Torture</td>
<td>CEM</td>
<td>-.65</td>
<td>.123</td>
<td>.000</td>
<td>1642</td>
</tr>
<tr>
<td>CEDAW</td>
<td>Pol. Rights</td>
<td>None</td>
<td>.513</td>
<td>.169</td>
<td>.002</td>
<td>1651</td>
</tr>
<tr>
<td>CEDAW</td>
<td>Pol. Rights</td>
<td>CEM</td>
<td>.306</td>
<td>.159</td>
<td>.056</td>
<td>1642</td>
</tr>
<tr>
<td>CEDAW</td>
<td>Soc. Rights</td>
<td>None</td>
<td>.428</td>
<td>.157</td>
<td>.006</td>
<td>1621</td>
</tr>
<tr>
<td>CEDAW</td>
<td>Soc. Rights</td>
<td>CEM</td>
<td>.15</td>
<td>.12</td>
<td>.211</td>
<td>1642</td>
</tr>
<tr>
<td>CEDAW</td>
<td>Econ. Rights</td>
<td>None</td>
<td>.209</td>
<td>.148</td>
<td>.159</td>
<td>1629</td>
</tr>
<tr>
<td>CEDAW</td>
<td>Econ. Rights</td>
<td>CEM</td>
<td>.039</td>
<td>.119</td>
<td>.747</td>
<td>1642</td>
</tr>
</tbody>
</table>
terms of the treaty. The same cannot be said for the results pertaining to the CAT.

**CAT**

Table 2 shows estimates from the CAT models, while Table 3 shows the SATT for the CAT. Examining the coefficients in the third and fourth rows of Table 4, we can see that ratification of the CAT is found to be associated with worse torture practices and this relationship actually becomes larger after matching. Again, the coefficient is not of as much interest to us as is the SATT. Since the measure of torture employed is ordinal (ranging from 0 to 2), the conditional probability of a state falling into a particular category was calculated for all three categories. Table 3 shows the average probability of falling into the first and last categories for ratifying states, the counterfactual probability of falling into these categories for ratifying states, and the SATT of the treaty, calculated as the difference between these probabilities. The first large-scale empirical investigation of the effects of human rights treaties on state practices, conducted by Oona Hathaway (2002) found that ratification of the Convention against Torture seemed to have a detrimental effect on torture practices, i.e. states who ratified the convention were more likely to practice torture. In light of this, finding a positive relationship between ratification of the CAT and the probability of torture should perhaps not be too surprising, however if one looks beyond coefficients to calculate substantive effects the results are startling.

The effect of the CAT among the states that have chosen to ratify it is quite large and negative (Recall that the torture variable is coded so that higher values indicate that torture occurs with lower frequency). Table 3 indicates that ratification of the CAT negatively affects the probability of falling into the highest category (torture does not occur), decreasing the probability of falling into this category by .16. Most disturbingly, ratification significantly raises the probability of falling into the worst category (torture occurs frequently). The CAT’s effect on the probability of torture occurring frequently is shown in the first column of Table 3. The table indicates that states that have ratified the CAT are, on average, .15 more likely to practice torture frequently than they would be had they not ratified the CAT.

**CEDAW**

Interestingly, the CEDAW seems to have had a positive, statistically significant impact on observance of at least one type of women’s right (political rights) among states that have ratified the convention. This is one of the first instances I am aware of where systematic empirical research has shown human rights treaties to have a beneficial effect on human rights practices.15 Table 2 shows the results from each of the women’s rights models, while the bottom portion of Table 3 shows the SATT of the CEDAW. We can see from Table 2 that the effect of the CEDAW on women’s political rights is positive and significant even after matching, though it becomes somewhat weaker. The results from the social rights models are markedly different. Though the treaty seems to have a positive impact on women’s social rights before matching, after the heterogeneity between treatment and control groups is reduced this effect disappears. The CEDAW was found to have no effect on observance of women’s economic rights before or after matching.

Turning now to substantive effects, Table 3 indicates that, among states that have ratified the CEDAW, the convention seems to have lead to a decrease of about .02 in the probability of being in the worst (i.e., the most abusive) category of the CIRI women’s political rights scale.16 Ratification also seems to have increased by about .05 the probability

---

15Simmons (2009) finds international human rights treaties to have a beneficial effect in some instances.

16Thought the SATT was calculated was calculated for all four categories of the political rights measure, it is shown for the first and last categories only.
that a state would be coded into the best category.\textsuperscript{17} While this change is quite modest it is an increase nevertheless, and it indicates that perhaps states do take their legal obligations under the human rights regime seriously in enough instances to produce a positive average impact.

\textbf{Conclusion}

One of the more interesting findings of this study is that ratification of the CAT is associated with worse torture practices than would be expected in the absence of ratification. While other scholars have found that the CAT is associated with the frequent use of torture (Hathaway 2002), its substantive impact has yet to be appreciated. Exactly why the CAT is associated with worse torture practices is not clear. Vreeland (2008) provides an explanation for why this would be the case amongst autocratic states, but why such a relationship would be found among all regime types is not obvious. One possible explanation is the “decoupling” effect posited by Hafner-Burton and Tsutsui (2005), whereby states ratify human rights treaties as a subterfuge to hide worsening practices. Such an effect may be present with respect to the ICCPR and even more so with respect to the CAT.

The other interesting finding here is the beneficial effect of the CEDAW on respect for women’s political rights. While the results pertaining to women’s social rights seem to tell the kind of story expected by those skeptical of the ability of international institutions to constrain state behavior (e.g., Downes, Rocke, and Barsoom 1996; von Stein 2005), the effect of the CEDAW on political rights remain even after the observed heterogeneity between ratifying and non-ratifying states is removed. One cannot help but wonder why these results are so markedly different from the results pertaining to the other treaties. A plausible answer lies in the nature of women’s rights in relation to threats to state sovereignty. Barring extreme cultural biases against women, improvements in women’s rights should be easy to effect relative to improvements in personal integrity rights. This is because women, as a subset of a state’s population, do not pose a threat to a leader’s position in power. Allowing women to enjoy equality with men in economic, social, and political matters should not pose a problem for most leaders.\textsuperscript{18} In contrast, alleviating personal integrity violations typically means curbing repression of dissidents, members of the population who do pose a threat (or are perceived to pose a threat) to the leader’s position in power. Leaders who are repressing dissidents are likely doing so to protect their tenures, whereas the reasons for systematic violations of women’s rights may be orthogonal to maintenance of authority. Thus it may be the case that the costs of complying with a women’s rights treaty are typically less than the costs of complying with treaties pertaining to personal integrity rights.

Whatever the reasons for compliance or non-compliance, the fact is that state behavior with respect to the international human rights regime exhibits quite a bit of variance. While the CEDAW shows promise for improving state behavior the CAT and, to a lesser extent, the CCPR seem to actually be associated with worse practices. The treaties examined here pertaining to personal integrity do not seem to have any beneficial effect on repressive behavior, possibly due to the connection between these rights and the maintenance of state authority. Scholars should consider the possibility that the costs of complying with human rights treaties depend on the kind of rights in question; where the authority of the state vis-a-vis domestic challenges is not impinged upon by the provisions of a treaty compliance may be less costly. This discussion implies that treaty-specific theory building may be needed. It also implies that there is some hope in using legal agreements to alter state practices if they pertain to issues that are less directly connected to state authority than personal integrity rights. Scholars should, in the future, examine the effects of treaties that fit this description, for example the Convention on the Rights of the Child. On the other hand, it may be the case that treaties designed to protect personal integrity rights have simply not had enough time to exert significant influence over state practices, particularly if the treaties’ main effect lies in their contribution to a normative discourse which alters individuals’ and states’ perceptions of appropriate behavior, an effect which is beyond the scope of the analysis presented here. Still, scholars should consider the possibility that, with respect to personal integrity rights, domestic circumstances create incentives for

\textsuperscript{17}The SATTs of the CEDAW for women’s social and economic rights were not calculated since the treaty coefficient was not statistically significant after matching.

\textsuperscript{18}Although in some places attempting to improve women’s rights may upset portions of the population that leaders depend on for support, we would expect this to be the case in few enough cases to make the average cost to leaders of improving women’s rights rather low.
rights violations that may be too strong to be mitigated by legal agreements.

Acknowledgments

The author thanks Jason Barabas, Andy Beger, Courtenay Conrad, Jaqueline Demeritt, Justin Esaray, Will Moore, Mark Souva, and Joe Young for helpful comments on earlier drafts.

Manuscript submitted 25 February 2009
Manuscript accepted for publication 11 February 2010

References


Daniel W. Hill, Jr. is a Ph.D. Candidate at The Florida State University, Tallahassee, FL 32301.