

## ***The Changing Standard of Accountability and the Positive Relationship between Human Rights Treaty Ratification and Compliance***

CHRISTOPHER J. FARISS\*

Researchers have puzzled over the finding that countries that ratify UN human rights treaties such as the Convention Against Torture are more likely to abuse human rights than non-ratifiers over time. This article presents evidence that the changing standard of accountability – the set of expectations that monitoring agencies use to hold states responsible for repressive actions – conceals real improvements to the level of respect for human rights in data derived from monitoring reports. Using a novel dataset that accounts for systematic changes to human rights reports, it is demonstrated that the ratification of human rights treaties is associated with higher levels of respect for human rights. This positive relationship is robust to a variety of measurement strategies and model specifications.

*Keywords:* human rights; treaty compliance; measurement; latent variable model; validity

*The normative appeal of international law is predicated upon the view that well-designed rules will – in general and on average – promote peace, stability and good governance.*

(Goodman and Jinks 2003, p. 171)

Why do states that ratify human rights treaties violate human rights more often than those countries that do not ratify such treaties? Researchers continue to puzzle over the empirical finding that countries that ratify the various instruments within the global human rights regime are more likely to abuse human rights than non-ratifiers over time.<sup>1</sup> In a recent paper, I have identified a possible answer to this question: the set of expectations used by monitoring agencies to hold states responsible for repressive actions has become increasingly strict over time.<sup>2</sup>

\* Jeffrey L. Hyde and Sharon D. Hyde and Political Science Board of Visitors Early Career Professor in Political Science, and Assistant Professor, Department of Political Science, Pennsylvania State University (e-mails: [cjf20@psu.edu](mailto:cjf20@psu.edu); [cjf0006@gmail.com](mailto:cjf0006@gmail.com)). The author would like to thank Daniel Berliner, Chad Clay, Charles Crabtree, Geoff Dancy, Jesse Driscoll, James Fowler, Danny Hill, Miles Kahler, David Lake, Milli Lake, Yon Lupu, Jamie Mayerfeld, Will Moore, Amanda Murdie, Michael Nelson, Keith Schnakenberg, Brice Semmens, Kathryn Sikkink, Reed Wood, and Thorin Wright for many helpful comments and suggestions. The code and data files necessary to implement the models in JAGS and R are available in Harvard Dataverse at: [doi:10.7910/DVN/TI77ZP](https://doi.org/10.7910/DVN/TI77ZP) and <https://dataverse.harvard.edu/dataverse/CJFariss>. Online appendices are available at <http://dx.doi.org/doi:10.1017/S000712341500054X>. This research was supported in part by the McCourtney Institute for Democracy Innovation Grant, Pennsylvania State University.

<sup>1</sup> The negative correlation found between ratification of various human rights treaties and respect for human rights (Hafner-Burton and Tsutsui 2005; Hafner-Burton and Tsutsui 2007; Hathaway 2002) has been criticized by some (Clark and Sikkink 2013; Goodman and Jinks 2003; Fariss 2014). However, this counterintuitive finding is generally assumed to be an empirical fact that is now taken for granted in the literature (Hafner-Burton 2013; Hafner-Burton and Ron 2009; Hafner-Burton, Tsutsui and Meyer 2008; Hollyer and Rosendorff 2011; Vreeland 2008) despite direct evidence to the contrary (Dancy and Sikkink 2012; Fariss 2014). This evidence is sometimes even overlooked or ignored (Hafner-Burton 2014).

<sup>2</sup> Fariss 2014.

Changes to this ‘standard of accountability’ mask real improvements to the level of respect for human rights in data derived from human rights monitoring reports, which implicitly incorporate these increasingly stringent assessments of state behaviors. Once this changing standard of accountability is taken into account using a new latent variable model of repression, the relationship between ratification of the UN Convention Against Torture and respect for human rights becomes positive.<sup>3</sup> This result suggests that (1) countries which respect human rights are more likely to ratify the UN Convention Against Torture in the first place, (2) the treaty has a causal effect on human rights protection once ratified, or even possibly (3) both. This new finding has broad implications for the human rights, international relations, and international law literatures, because it suggests that the international human rights regime is not simply cover for human rights abusers but is instead associated with respect for human rights and improvements in state behaviors over time. But what about the relationship between respect for human rights and the many other international human rights conventions that are a part of the global human rights regime?

In this article, I demonstrate that there are systematic differences in the relationship between the level of respect for human rights – when the changing standard of accountability is and is not accounted for – and several widely studied international human rights treaties (measured in several different ways). Specifically, correlations that were once found to be negative are actually positive. I demonstrate that these new positive relationships hold generally for a set of human rights treaties, which are measured using a standard additive approach commonly used in the treaty compliance literature to capture the level of embeddedness of a state within the international human rights regime over time.<sup>4</sup> I also introduce a new measure of the global human rights regime using a dynamic latent variable model similar to the one developed by Martin and Quinn to model the ideology of Supreme Court Justices over time.<sup>5</sup> Overall, when changes in the standards used to assess state abuse are taken into account, the negative relationship between respect for human rights and ratification of human rights treaties is reversed across several different human rights treaty variables. The results suggest that the ‘normative appeal of international law’ may not be in doubt as earlier empirical research suggests.<sup>6</sup>

In the remainder of this article, I first describe the disagreement that exists within the literature on international treaty compliance and how this relates to the changing standard of accountability. The acknowledgment and assessment of the standard of accountability has major implications for this debate. Next, I describe the existing measurement strategies used to assess human rights behaviors and treaty compliance and then demonstrate how the new human rights data changes existing relationships with both old and new measurements of treaty ratification over time. Finally, I conclude with suggestions for future research.

#### HUMAN RIGHTS TREATY COMPLIANCE AND THE CHANGING STANDARD OF ACCOUNTABILITY

In this section, I describe the disagreement between scholars on state compliance with international human rights treaties. I then describe how changes to the standard of accountability over time help to explain the anomalous empirical findings reported in earlier studies, which catalyzed this debate.

Understanding the relationship between respect for human rights and the effectiveness of UN human rights treaties is important for the human rights, international relations, and international

<sup>3</sup> Fariss (2014, pp. 311–13).

<sup>4</sup> See, for example, Hafner-Burton and Tsutsui (2005).

<sup>5</sup> Martin and Quinn 2002.

<sup>6</sup> Goodman and Jinks (2003, p. 171).

law literatures because there are two divergent arguments about treaty effectiveness. Some authors argue that treaty ratification constrains states by creating costs for noncompliance, which modifies state behaviors.<sup>7</sup> Alternatively, treaty ratifications only occur in cases in which the regime would have complied with the provisions of a treaty regardless of whether or not the treaty was ever ratified. Thus, treaties have no effect on the codified behaviors within the specific treaty, such as the level of cooperation,<sup>8</sup> or human rights compliance.<sup>9</sup> According to this group of scholars, treaty ratification is simply a reflection of the preferences of the political leader. The theory and data presented in this article cast considerable doubt on this second interpretation.

As described above, there is considerable debate about the reasons for the negative correlations found between respect for human rights and treaty compliance reported in earlier research.<sup>10</sup> Using more sophisticated research design strategies that provide more plausible evidence of causal effects, scholars have found evidence for positive relationships between treaty ratification and human rights compliance under certain conditions.<sup>11</sup> For example, ratification of international treaties is associated with improvement in women's protections and select civil and political rights under certain conditions.<sup>12</sup> Even more recent evidence suggests that ratification of the Convention Against Torture actually reduces torture in cases in which leaders are secure in their jobs<sup>13</sup> or when many legislative veto players are present and constrain political leaders based on rules written into law.<sup>14</sup>

Unfortunately, with the exception of my recent article, all of the published empirical assessments of these opposing viewpoints<sup>15</sup> have made use of data measuring repression derived from the same human rights reports.<sup>16</sup> The data used in these studies do not account for the changing standard of accountability, which confounds the relationship between treaty compliance and respect for human rights.

As I demonstrate in this article, the acknowledgment and assessment of the standard of accountability has major implications for the treaty compliance debate because it identifies changes to human rights reporting that correspond in time with the increasing embeddedness of countries within the global human rights regime. Importantly, the empirical evidence presented in this article, based on new human rights estimates that account for the changing standard of accountability, bolsters the positive statistical associations for the effectiveness of the human rights regime on human rights behaviors reported in some of these more recently published studies.<sup>17</sup> This occurs because these previously reported positive effects were discovered using biased data that did not account for changes to the standard of accountability over time. That is, the substantive effects reported in these published studies are likely to be stronger than

<sup>7</sup> Morrow 2007; Simmons 2000; Simmons and Hopkins 2005.

<sup>8</sup> Downs, Rocke and Barsoom 1996; Von Stein 2005.

<sup>9</sup> Hathaway 2002; Hafner-Burton and Tsutsui 2005; Hafner-Burton and Tsutsui 2007.

<sup>10</sup> Hathaway 2002; Hafner-Burton and Ron 2009; Hafner-Burton and Tsutsui 2005; Hafner-Burton and Tsutsui 2007.

<sup>11</sup> Conrad 2014; Conrad and Ritter 2013; Dancy and Sikkink 2012; Landman 2005; Lupu 2013a; Lupu 2015; Simmons 2009; Smith-Cannoy 2012.

<sup>12</sup> Lupu 2013a; Simmons 2009.

<sup>13</sup> Conrad and Ritter 2013.

<sup>14</sup> Lupu 2015.

<sup>15</sup> See Fariss 2014; Conrad 2014; Conrad and Ritter 2013; Dancy and Sikkink 2012; Hathaway 2002; Hafner-Burton and Tsutsui 2005; Hafner-Burton and Tsutsui 2007; Hill 2010; Keith 1999; Landman 2005; Lupu 2013a; Lupu 2015; Neumayer 2005; Simmons 2009.

<sup>16</sup> Cingranelli and Richards 1999; Cingranelli, Richards and Clay 2015a; Cingranelli, Richards and Clay 2015b; Gibney, Cornett and Wood 2012; Hathaway 2002.

<sup>17</sup> Conrad and Ritter 2013; Dancy and Sikkink 2012; Landman 2005; Lupu 2013a; Lupu 2015; Simmons 2009.

previously thought. Finally, the results reported in this article directly contradict the negative correlations reported in earlier studies<sup>18</sup> and cast considerable doubt on studies that begin with this negative correlation as a puzzle needing to be explained.<sup>19</sup>

Why do changes to the standard of acceptability confound the relationship between treaty ratification and compliance? In my previous article I have shown that the standard of accountability – ‘the set of expectations that monitoring agencies use to hold states responsible for repressive actions’ – has changed due to a combination of three mechanisms or tactics: (1) information, (2) access, and (3) classification.<sup>20</sup> Over time, observers and activists update these tactics in order to reveal repressive practices to the international community, understand why those practices occurred in the first place, and eventually change those practices for the better. Over the same period of time, however, state officials have signed and ratified an increasing number of UN human rights treaties, thus becoming increasingly embedded in the international human rights regime. To untangle the relationship between treaty ratification and state compliance then, it is also necessary to understand the process by which human rights reporting changes over time.

To understand how human rights reporting changes over time, first consider the amount and quality of *information* available to monitoring agencies with which to document the repressive behaviors of states. The standard of accountability changes because of improvements in the quality and increases in the quantity of information. These changes lead to more accurate assessments of the conditions in countries each year. Thus, the global pattern of human rights may appear stagnant or possibly even worse over time because of an increasing amount of information about how states respect or abuse human rights.<sup>21</sup> Over time, monitoring agencies are looking harder for abuse with more and better information at the same time that state officials are signing and ratifying an increasing number of UN human rights treaties.

Second and relatedly, consider the level of *access* to countries by non-governmental organizations (NGOs) like Amnesty International and Human Rights Watch. These organizations, and others like them, seek to gather, corroborate, and publish accurate information about allegations of human rights abuses. The ability to accomplish this goal has increased as these organizations grow and cooperate with one another. For example, Amnesty International and the US State Department gather allegations of human rights abuse from the reports of other NGOs that collect, publish, and share the information gathered from local or regional contexts.<sup>22</sup> Overall, the standard of accountability changes ‘as access to government documents, witnesses, victims, prisons sites, and other areas’, essential for assessing state compliance with international human rights law, increases.<sup>23</sup> Over time, monitoring agencies are looking in more places for abuse at the same time as state officials are signing and ratifying an increasing number of UN human rights treaties.

Third, consider how the *classification* of different types of abuse changes over time. That is, changes to the subjective views of what constitutes a good human rights record by the agencies and activists that monitor state behaviors are conditioned by the overall level of abuse or respect

<sup>18</sup> Hafner-Burton and Tsutsui 2005; Hafner-Burton and Tsutsui 2007; Hathaway 2002.

<sup>19</sup> Hafner-Burton 2013; Hafner-Burton, Tsutsui and Meyer 2008; Hollyer and Rosendorff 2011; Posner 2014; Vreeland 2008.

<sup>20</sup> Fariss (2014, p. 299).

<sup>21</sup> Keck and Sikkink (1998) and Clark and Sikkink (2013) argue that this change occurs because of an ‘information paradox’ or ‘human rights information paradox’. ‘The paradox occurs when an increase in information leads to difficulties in assessing the efficacy of an advocacy campaign over time because of the very success in collecting and aggregating information about the use of the repressive policy in the first place’ (Fariss 2014, p. 299).

<sup>22</sup> Hill, Moore and Mukherjee 2013; Hopgood 2006; Keck and Sikkink 1998; Korey 2001; Murdie and Bhasin 2011; Murdie and Davis 2012; Wong 2012.

<sup>23</sup> Fariss (2014, p. 300).

for human rights across countries in the international system. Today, what an observer at Amnesty International or the State Department considers to be a ‘good’ record of human rights is more stringent when compared to records from earlier years because the global average of respect for human rights has also improved. These changes occur as observers pressure governments to institute new reforms, even after other reforms have been implemented to reduce more egregious rights violations such as extra-judicial killings and disappearances. As I have argued elsewhere, ‘monitoring agencies are increasingly sensitive to the various kinds of ill-treatment that previously fell short of abuse but that still constitute violations of human rights’.<sup>24</sup> Moreover, there is even evidence from human rights case law of a rising standard of acceptable treatment, in which more state behaviors are classified as torture.<sup>25</sup> Again, monitoring agencies are classifying more acts as abuse at the same time as state officials are signing and ratifying an increasing number of UN human rights treaties over time.

Thus the relationship between treaty compliance and respect for human rights is confounded by the changing standard of accountability. What this all means for scholars of human rights, international relations, and international law, is that more recent versions of annual human rights reports represent a broader and more detailed view of the human rights practices than reports published in previous years.<sup>26</sup> As Sikkink notes, these organizations ‘have expanded their focus over time from a narrow concentration on direct government responsibility for the death, disappearance, and imprisonment of political opponents to a wider range of rights, including the right of people to be free from police brutality and the excessive use of lethal force’.<sup>27</sup> By allowing the standard of accountability to vary with time, ‘a new picture emerges of improving physical integrity practices over time (1949–2010) since hitting a low point in the mid-1970s’.<sup>28</sup> Though this new view should be welcomed by scholars and activists alike, these changes call into question empirical analyses of data coded from these reports because again, the changing standard of accountability confounds the relationship between treaty compliance and respect for human rights when it is not accounted for by existing measures of human rights practices. Fortunately, new data that corrects for the changing standard of accountability are now publicly available.<sup>29</sup>

In the remaining sections of this article, I use new, publicly available human rights data to address this debate. To preview the results, I find that unobserved changes to the standard of accountability explain why correlations between all reports-based human rights variables and

<sup>24</sup> Fariss (2014, p. 300).

<sup>25</sup> *The European Court of Human Rights, in Selmouni v. France* (1999), ‘consider certain acts which were classified in the past as inhuman and degrading treatment as opposed to torture could be classified differently in future’. That is, acts by state agents that might have previously been classified within the less severe category of ill-treatment and degrading punishment might now be classified as torture. The court states further ‘that the increasingly high standard being required in the area of the protection of human rights and fundamental liberties correspondingly and inevitably requires greater firmness in assessing breaches of the fundamental values of democratic societies’. See *Selmouni v. France*, 25803/94, Council of Europe: European Court of Human Rights, 28 July 1999, available at: <http://www.unhcr.org/refworld/docid/3ae6b70210.html>.

<sup>26</sup> The primary documents I reference are the *Country Reports on Human Rights Practices* published annually by the US State Department and *The State of the World’s Human Rights* report published annually by Amnesty International.

<sup>27</sup> Sikkink (2011, p. 159).

<sup>28</sup> Fariss (2014, p. 314).

<sup>29</sup> The validity of the reports-based human rights variables only comes into question when researchers make the conceptual decision that the values of those variables represent the ‘true’ level of human rights abuse instead of the reported level of human rights abuse. As I discuss (2014, p. 316), this is an important theoretical distinction, which is often overlooked when the original PTS, CIRI, and Hathaway variables (Cingranelli and Richards 1999; Cingranelli, Richards and Clay 2015a; Cingranelli, Richards and Clay 2015b; Gibney, Cornett and Wood 2012; Hathaway 2002) are presented as measurements of *abuse* instead of *reported abuse*.

various human rights treaty variables are negative. When the new, corrected data are used in place of uncorrected data, the correlations between respect for human rights (i.e., compliance) and treaty ratification variables become positive. Like the relationship between the new human rights variable and ratification of the Convention Against Torture, these new results suggest (1) that countries which respect human rights are more likely to ratify UN human rights treaties in the first place, (2) that these treaties have a causal effect on human rights protection once ratified, or possibly even (3) both.

#### LATENT VARIABLES MODELS FOR UNOBSERVABLE CONSTRUCTS

In this section, I present information about latent variable models in three subsections. First, I present a general introduction to latent variables models and an overview of applications from various subfields in political science and related fields. Second, I present an applied version of the model, which is a dynamic binary latent variable mode, which measures the level of ‘embeddedness’ of a country within the global human rights regime over time.<sup>30</sup> In this subsection, I describe how I use this model to develop new latent treaty ratification estimates, which I contrast with two alternative additive treaty ratification scales and a variable that measures the proportion of ratified treaties. Third, I describe a similar latent variable model and the publicly available latent human rights estimates that I have developed, which extends the dynamic ordinal latent variable model introduced by Schnakenberg and Fariss.<sup>31</sup> Additional information about these models and the variables used to estimate them are available in the supplementary appendix.

#### *An Introduction to Latent Variable Models in Political Science*

Examples of latent variable or Item Repose Theory (IRT) models are now quite ubiquitous in the social sciences and political science, especially in the study of legislative ideology in the United States.<sup>32</sup> Scholars have also used these models in the study of international relations<sup>33</sup> and comparative politics generally<sup>34</sup> and in the study of human rights in particular.<sup>35</sup> Dynamic versions of these models are also increasingly common. The DW-NOMINATE procedure is a dynamic version of W-NOMINATE, which estimates the ideal points of members of a legislature as a function of ideal points from the previous time period.<sup>36</sup> Martin and Quinn introduced a Bayesian dynamic IRT model to estimate ideal points in the United States Supreme Court based on binary decision data and model the temporal dependence in these data by specifying a prior for each value of the latent variable centered at the estimated latent variable from the same unit in the previous time period.<sup>37</sup> Schnakenberg and Fariss build on this insight in order to extend the ordinal IRT model introduced by Treier and Jackman.<sup>38</sup> More recently,

<sup>30</sup> The dynamic binary latent variable model is a dynamic version of a binary item response theory model. I describe the development and use of this model in the subsections below.

<sup>31</sup> The dynamic ordinal latent variable model developed by Schnakenberg and Fariss (2014) and extended by Fariss (2014) is a dynamic version of a static ordinal item response theory model or graded-response model (Albert and Johnson 1999) first introduced to political science by Treier and Jackman (2008).

<sup>32</sup> Barbera 2015; Bonica 2012; Caughey and Warshaw 2015; Clinton, Jackman and Rivers 2004; Poole and Rosenthal 1997. See also Bond and Messing (2015).

<sup>33</sup> Lupu 2013b; Voeten 2000.

<sup>34</sup> Pemstein, Meserve and Melton 2010; Treier and Jackman 2008.

<sup>35</sup> Fariss 2014; Schnakenberg and Fariss 2014.

<sup>36</sup> Poole and Rosenthal 1997.

<sup>37</sup> Martin and Quinn 2002.

<sup>38</sup> Schnakenberg and Fariss 2014; Treier and Jackman 2008.

Fariss extends this model even further by allowing some of the item-difficulty parameters (the threshold parameters on the ordered human rights variables) to vary over time instead of being held constant.<sup>39</sup> That is, the baseline probability of a human rights variable being coded at a specific value is allowed to change over time.<sup>40</sup>

The setup for the basic version of these models is quite intuitive. For the most part, the data of interest to scholars in international relations and comparative politics are made up of country–year observations indexed by  $i$  and  $t$ :  $i = 1, \dots, N$  indexes countries and  $t = 1, \dots, T$  indexes year. Of course, other grouping structures for different units are possible (e.g., students in schools, or survey respondent across counties, states, or even countries). The latent variable  $\theta$  (theta) is inferred from the modeled relationship between it and the observed manifest data or items. The term ‘item’ was first used by researchers developing educational tests and refers to the correct and incorrect responses to test questions on intelligence assessments.<sup>41</sup> The observed items can take on any numeric value or even be missing. Data can be ratio (continuous with a true 0), interval (continuous without a true 0), categorical (ordered categories), nominal (unordered categories), or a mix of these types. For example, the data can contain continuous indicators,<sup>42</sup> event counts,<sup>43</sup> values from categorized documents,<sup>44</sup> nominal question responses of unknown order,<sup>45</sup> or again, some mix of different data types.<sup>46</sup>

A simple way to begin constructing a latent variable model begins with a linear regression equation (without reference to the country–year subscripts  $i$  to  $t$  or any other grouping referent):

$$y = \alpha + \beta x + \varepsilon \quad (1)$$

This model is similar to a factor analytic model but the label for the model is not an important point.<sup>47</sup>  $\alpha$  represents the intercept of the model,  $\beta$  is the slope, and  $\varepsilon$  is the error term. Below, I drop  $\varepsilon$  and change the equals sign ( $=$ ) to a tilde sign ( $\sim$ ), which simply denotes a stochastic relationship between the dependent variable  $y$  and the independent variable  $x$ ,

$$y \sim \alpha + \beta x. \quad (2)$$

Equations 1 and 2 are equivalent, though I have not specified the structure of the error process for either equation. The structure of the error is not important at this point. When  $x$  is observed, both of these equations are simple linear models and are easy to estimate. The model becomes noticeably more complicated when  $x$  must also be estimated. Below, I change  $x$  to the Greek symbol theta ( $\theta$ ) to denote that it is also a parameter to be estimated:

$$y \sim \alpha + \beta \theta. \quad (3)$$

In Equation 3,  $x$  is replaced by  $\theta$ , which is the estimated ‘true’ or latent variable to be estimated. The theoretical interpretation of the model is that  $y$  is caused by the ‘true’ level of the latent variable  $\theta$ . Stated differently;  $y$  is an observable outcome, which arises from the theoretical concept of interest. Importantly, the observation of  $y$  does not need to be perfect, especially when the researcher has an understanding of the biases related to the data generating

<sup>39</sup> Fariss 2014.

<sup>40</sup> This is how the changing standard of accountability is incorporated into the latent variable model of human rights.

<sup>41</sup> Lord 1980; Lord and Novick 1968; Rasch 1980. See Borsboom (2005) and Jackman (2008) for accounts of the development of this literature.

<sup>42</sup> Quinn 2004.

<sup>43</sup> Fariss 2013.

<sup>44</sup> Schnakenberg and Fariss 2014.

<sup>45</sup> Jessee 2015.

<sup>46</sup> Fariss 2014; Quinn 2004.

<sup>47</sup> In a factor analysis model,  $\alpha$  is set to 0 and  $\beta$  is analogous to a factor loading.

process for a given observed item  $y$ . The more information obtained about the political process by which the observation of  $y$  occurs, the more closely it can be related to the model of the latent variable or theoretical concept of interest.<sup>48</sup>

Again,  $\alpha$  is the intercept or for categorical data, the baseline probability of observing a given category of  $y$ . In many latent variable applications,  $\alpha$  represents the ‘difficulty’ in being coded at a certain level of the latent variable.  $\beta$  is the slope or the strength of the relationship between the ‘true’ level of the latent variable  $\theta$  and the value of the observed variable  $y$ . In many latent variable applications,  $\beta$  represents the ability of a test or model to ‘discriminate’ between different values along the latent variable or trait (e.g., intelligence, ideology, treaty embeddedness, repression). This is a common specification for the relationship between the observed manifest variables and the latent variables.

Importantly, I wish to reiterate that there is no way to estimate a latent variable based on a single indicator.<sup>49</sup> If only one test score exists, then the only value to estimate the latent variable  $\theta$  is the single test score itself. In this case  $\alpha$  and  $\beta$  are irrelevant parameters that equal 0 and 1 respectively. However, it is possible to begin to approximate an estimate of  $\theta$  with more observed information about the concept of interest. Each of the  $j = 1, \dots, J$  indicators are for the various observed indicators that are observable manifestations of the theoretical concept of interest (e.g., intelligence, ideology, treaty embeddedness, respect for human rights). As the number of  $j$  items  $y_j$  increases, the precision of the estimates of  $\theta$  also increases. As Jackman points out, a researcher with only one indicator of a latent construct is unable to determine how much variation in the indicator is due to measurement error as opposed to other forms of variation in the latent construct.<sup>50</sup> Adding more observed information into the latent variable model allows for a more precise estimate of  $\theta$ , and can be formalized for continuous variables with the following structural equation:

$$Y_1 \sim \alpha_1 + \beta_1 \theta$$

$$Y_2 \sim \alpha_2 + \beta_2 \theta$$

$$\cdot \quad \cdot \quad \cdot$$

$$\cdot \quad \cdot \quad \cdot$$

$$\cdot \quad \cdot \quad \cdot$$

$$Y_j \sim \alpha_j + \beta_j \theta.$$

Note that  $\theta$  is the same for every equation, though the model parameters (here  $\alpha$  and  $\beta$ ) change for each equation, or more generally the function  $f(\cdot)$ . This identification strategy allows for the estimation of the latent variable and, in more complex versions of the model, the level of

<sup>48</sup> As an important aside, the Bayesian perspective on the relationship between data and model parameters provides a foundation for using the often flawed and incomplete data with which scholars of human rights, repression, and contentious politics must work in order to generate new insights about these important concepts. This perspective is useful because it shifts the burden of validity away from the primary source documentation and raw data to the model parameters that bind these diverse pieces of information together. This perspective emerges when considering the sources of our uncertainty, which are fundamentally part of any validity assessment. ‘With conventional statistics, the only uncertainty admitted to the analysis is sampling uncertainty. The Bayesian approach offers guidance for dealing with myriad sources of uncertainty faced by applied researchers in real analyses’ (Western 1999, p. 20, as quoted by Gill 2008, p. 28).

<sup>49</sup> Jackman 2008; Schnakenberg and Fariss 2014.

<sup>50</sup> Jackman 2008.

uncertainty about the parameter estimate. The level of uncertainty is formalized by the relative agreement or disagreement between scores across multiple indicators. The more disagreement across observed variables, the greater the uncertainty in the resulting latent variable estimates for the individual subject.

What is left for the researchers interested in estimating an unobservable concept, is the choice in what observed variables (manifest variables or items) to include in the latent variable model, the level of organization of the grouping information or subscripts (i.e., how one unit relates to another), and the function that relates the observed variable to the latent variable ( $\alpha$  and  $\beta$  are part of the linear function assumed above).<sup>51</sup> These are each choices that should be driven by a theory of the concept of interest and then assessed using the tools of construct validity, which allow for both the assessment of the latent variable as it relates back to the theoretical concept (translation validity or how closely the estimate of  $\theta$  relates to the theoretical concept) and as it relates to alternative representations of the latent variable (criterion-related validity or how the estimate of  $\theta$  relates to other variables that measure the same concept).<sup>52</sup>

This modeling approach provides a foundation for using the often flawed and incomplete data with which scholars of repression, human rights, and contentious politics must work in order to generate new insights about these important concepts. This perspective is useful because it shifts the burden of validity from the primary source documentation and raw data to the model parameters that bind these diverse pieces of information together.

### *Measuring the Treaty Ratification and Embeddedness of States within the International Human Rights Regime*

Here, I use a dynamic binary item-response theory (IRT) model, similar to the model introduced by Martin and Quinn and quite similar to the general latent variable model described above. The model provides estimates for the level of ‘embeddedness’ of a country within the global human rights regime over time. Below, I compare this new latent variable estimate of ‘embeddedness’ with two alternative additive scales and a proportion variable each of which measure the same concept. The additive scales are simply sums of the number of treaties a country has ratified in a given year. The proportion variable is the ratio of the total number of ratified treaties for a given country out of the total number of treaties available for ratification in each year.

The dynamic binary IRT model is quite intuitive. As described in the model above, the data are made up of country–year observations indexed by  $i$  and  $t$ ;  $i = 1, \dots, N$  indexes countries and  $t = 1, \dots, T$  indexes year. Each of the  $j = 1, \dots, J$  indicators are for the various global treaties and optional protocols that cover a variety of human rights issues. Table 1 contains a listing of commonly studied human rights treaties that can be ratified by any country in the international system. The data can take on a value of ‘NA’, ‘0’, or ‘1’. A country–year observation is coded 1 for the year of ratification of a specific treaty and also 1 in every following year. The NAs are important because each of the treaties opens for ratification at different points in time. Thus, treaties with more ‘missingness’ will be less informative than those with less missingness over time. The uncertainty for each latent treaty variable for a given country–year is based on the number of treaties that could potentially have been ratified in a given year. As Table 1 shows,

<sup>51</sup> For models that do not assume a linear relationship between observed indicators and the latent variable, define a function  $f(\cdot)$  that relates the observed variables to the latent variable:  $y \sim f_j(\theta)$ . Latent variable models with binary data often use logistic regression equations to link the observed binary responses to the latent variable:  $y_j \sim \text{logit}^{-1}(\theta)$ .

<sup>52</sup> For more on construct validity and related research design issues, see Adcock and Collier (2001); Shadish (2010); Shadish, Cook and Campbell (2001) and Trochim and Donnelly (2008).

TABLE 1 *Global International Human Rights Instruments*

Treaty name	Signed*	In force†
Convention on the Prevention and Punishment of the Crime of Genocide	1948	1951
International Convention on the Elimination of All Forms of Racial Discrimination	1965	1969
International Covenant on Civil and Political Rights	1966	1976
International Covenant on Civil and Political Rights Optional Protocol	1966	1976
International Covenant on Economic, Social and Cultural Rights	1966	1976
International Convention on the Suppression and Punishment of the Crime of Apartheid	1971	1973
Convention on the Elimination of all Forms of Discrimination against Women	1979	1981
Convention Against Torture	1984	1985
Convention Against Torture Article 21 (no reservation)	1984	1985
Convention Against Torture Article 22 (no reservation)	1984	1985
Convention on the Rights of the Child	1989	1990
International Covenant on Civil and Political Rights Optional Protocol 2	1989	1991
Convention on the Protection of the Rights of All Migrant Workers and Members of Their Families	1990	2003
International Criminal Court	1998	2002
Convention on the Elimination of all Forms of Discrimination against Women Optional Protocol	1999	2000
Convention on the Rights of the Child Optional Protocol 1	2000	2002
Convention on the Rights of the Child Optional Protocol 2	2000	2002
Convention Against Torture Optional Protocol	2002	2006
Convention on the Rights of Persons with Disabilities	2006	2007
Convention on the Rights of Persons with Disabilities Optional Protocol	2006	2008
International Convention for the Protection of All Persons from Enforced Disappearance	2006	2010
International Covenant on Economic, Social and Cultural Rights Optional Protocol	2008	2013
Convention on the Rights of the Child Optional Protocol 3	2011	–

Sources: (1): University of Minnesota Human Rights Library, <http://www1.umn.edu/humanrts/>

(2): United Nations Treaty Collections, <http://treaties.un.org/>

\*The 'Signed' column refers to the year the treaty is opened for signature.

†The 'Force' column refers to the year the treaty enters into force.

fewer treaties exist in earlier decades. The latent variable therefore represents both the level of 'embeddedness' of a country within the global human rights regime and the propensity to ratify new treaties as they open for ratification over time. The uncertainty for each latent treaty variable for a given country–year is based on the number of treaties that could potentially have been ratified in a given year. The error terms  $\varepsilon_{ijt}$  are independently drawn from a logistic distribution, where  $F(\cdot)$  denotes the logistic cumulative distribution function. Each of the  $j$  treaty variables are binary  $y_{ijt}$  so the probability distribution is:

$$P[y_{ijt} = 1] = F(\alpha_j + \theta_{ij}\beta_j) \quad (4)$$

Assuming local independence of responses across units, the likelihood function for  $\beta$ ,  $\alpha$ , and  $\theta$  given the data is:

$$L(\beta, \alpha, \theta | y) = \prod_{i=1}^N \prod_{t=1}^T \prod_{j=1}^J \left[ F(\alpha_j + \theta_{ij}\beta_j)^{y_{ijt}} * (1 - F(\alpha_j + \theta_{ij}\beta_j))^{(1-y_{ijt})} \right] \quad (5)$$

Equation 4 refers to the probability of observing  $y_{ijt} = 1$ , and one minus this equation refers to the probability of observing  $y_{ijt} = 0$ . The likelihood Equation 5 refers to the probability of the observed value in the data  $y_{ijt}$ . I estimate the model using independent standard normal priors on the latent treaty variable  $\theta_{it}$ . In other words:

$$\theta_{it} \sim N(0, 1)$$

for all  $i$  when  $t = 1$ . The standard normal prior when  $t > 1$  is centered around the latent variable estimate from the previous year such that:

$$\theta_{it} \sim N(\theta_{it-1}, \sigma)$$

This method for incorporating dynamics was implemented in the context of a dichotomous IRT by Martin and Quinn.<sup>53</sup> One difference between this model and the Martin and Quinn model is that  $\sigma$  is estimated instead of specifying it *a priori*.<sup>54</sup>

The prior for variance  $\sigma$  is modeled as  $U(0, 1)$ . This reflects prior knowledge that the between-country variation in will be much higher on average than the average within-country variance.<sup>55</sup> Slightly informative gamma priors *Gamma*(4, 3) were specified for the  $\beta$  parameters. The  $\alpha$  parameters are given  $N(0, 4)$  priors again for all of the  $j$  treaties.<sup>56</sup>

The prior distributions for each of the model parameters help to provide a theoretical rationale for linking the latent variable  $\theta_{it}$  to the observed treaty variables  $y_j$ . The normal distribution for the latent variable  $\theta_{it}$  arranges the country–year latent variable estimates under a normal density curve with mean 0 and standard deviation of 1. The mean from year to year need not be 0 and will depend on the constellation of data points for each item but the global mean for all the country–years in the model should be approximately 0. These priors are not highly informative but they do impose an overall density on the distribution of the country–year latent variable estimates. The choice in density functions, in this case a normal distribution, is up to the researcher and can be changed based on theory or new empirical information. Again though, it is important then to evaluate the output of the model using the tools of construct validity.<sup>57</sup> Schnakenberg and Fariss provide several different examples of empirical comparisons that are useful for assessing the validity of the estimate from different latent variable models.<sup>58</sup>

The model assumes that any two item responses are independent conditional on the latent variable. This means that two item-responses are only related because of the fact that they are each an observable outcome of the same latent trait. In this case, the latent trait is the level of embeddedness of a country in the global human rights regime. There are three relevant local independence assumptions: (1) local independence of different indicators within the same country–year, (2) local independence of indicators across countries within years, and (3) local

<sup>53</sup> Martin and Quinn 2002.

<sup>54</sup> Schnakenberg and Fariss (2014) introduced this feature to their dynamic model.

<sup>55</sup> The estimates of  $\sigma$  from the posterior of the converged model illustrate that the distribution is nowhere near 1, so the truncation decision was not important.

<sup>56</sup> As is generally true of item-response models, the likelihood function is not identified. In particular, IRT models suffer from ‘invariance to reflection’, which means that multiplying all of the parameters by  $-1$  would have no effect on the likelihood function. Lack of identification is problematic in maximum likelihood models but is not a problem for Bayesian approaches. The problem of invariance to rotation motivated the choice to give the  $b$  parameters strictly positive priors. For more information on identification problems in IRT models, see Jackman (2009). The latent treaty model is estimated with two MCMC chains, which are run for 100,000 iterations using JAGS (Plummer 2010) on the Gordon Supercomputer (Sinkovits et al. 2011). The first 50,000 iterations were thrown away as burn-in and the rest were used for inference. Diagnostics all suggest convergence (Gelman and Rubin 1992; Geweke 1992; Heidelberger and Welch 1981; Heidelberger and Welch 1983).

<sup>57</sup> Adcock and Collier 2001; Trochim and Donnelly 2008.

<sup>58</sup> Schnakenberg and Fariss 2014.

independence of indicators across years within countries. The third assumption is relaxed by incorporating temporal information into prior beliefs about the latent treaty variable.

Ratification is not a uniform process. Thus, a useful feature of the model is that it can also account for reservations, understandings, and declarations accompanying ratification of a treaty. States can tailor their ratification of treaties, and these adjustments have legal implications. For example, states may attach reservations that cancel application of particular provisions, or they may adopt optional provisions that increase their obligations under a treaty. These adjustments are typically referred to as reservations, understandings, or declarations.<sup>59</sup>

I consider two important reservations that a state might make about the Convention of Torture in this article. Accounting for these different types of provisions presents a difficult modeling problem. As is mentioned in Dancy and Sikkink:

It should be noted here that ratification is not a uniform process, and it may vary from state to state. For example, states may issue declarations of reservation upon ratification, which may express an unwillingness to accept in full the provisions of the treaty. However, we do not distinguish between types of ratification for two reasons: first, it is not commonly done in the quantitative literature, on which we are building; and second, it is quite difficult to generate a coding scheme that accounts for reservations. The reason is that any reduction in the score assigned to a reserving country would be arbitrary, given that reservations are not always similar. Future research should work to generate new ways of dealing with this issue.<sup>60</sup>

The dynamic binary IRT model can systematically account for such differences using a simply binary coding scheme for the observed variables or items. For this model, I have elected to include two binary variables, each coded 1 if a reservation, understanding, or declaration does not exist for Article 21 and Article 22, conditional on ratification of the Convention Against Torture. If this convention is ratified but with a reservation, understanding, or declaration to an article, then the binary indicators are coded 0. Otherwise the variable is coded as NA. Note again that the NA coding has substantive meaning in the context of the IRT model. Each country that ratifies the Torture Convention can choose whether the Committee against Torture may hear complaints brought against it by another state, under the optional mechanism in Article 21, or by an individual, under the optional mechanism in Article 22. It is not uncommon for states to ratify the Convention Against Torture with a reservation or declaration about these two Articles. Specifically, a country declares itself willing to recognize the Committee's competence to hear inter-state complaints under Article 21 or individual complaints under Article 22. Overall, the model provides a principled way to incorporate this information along with the other data about treaty ratification generally.

I should reiterate that there is no model-free way to estimate a latent variable.<sup>61</sup> An additive scale approach, typical for modeling treaty ratification behaviors,<sup>62</sup> is a *model* assuming equally weighted indicators and no error. The new latent treaty variable provides an alternative to such a model by estimating the item-weights and the uncertainty of the estimates. I compare this new

<sup>59</sup> However, according to Roth (2001, p. 891, n. 2), 'A qualification attached to the instrument of ratification is a "Reservation", regardless of designation, wherever it manifests an intent to withhold consent to a treaty obligation. The terms "Understanding" and "Declaration" have no specified significance in international law; thus, whether qualifications under these headings amount to reservations is a question of intent'. Mayerfeld (2016, chap. 5) discusses this issue among many others in the draft of his forthcoming book *The Effectiveness of International Human Rights Law*. I offer special thanks to Jamie Mayerfeld for helping me to clarify these important legal distinctions.

<sup>60</sup> Dancy and Sikkink (2012, p. 771, n. 51).

<sup>61</sup> See the detailed discussion of the issue of measurement models more generally in both Fariss and Schnakenberg (2014) and Schnakenberg and Fariss (2014).

<sup>62</sup> Hafner-Burton and Tsutsui 2005.

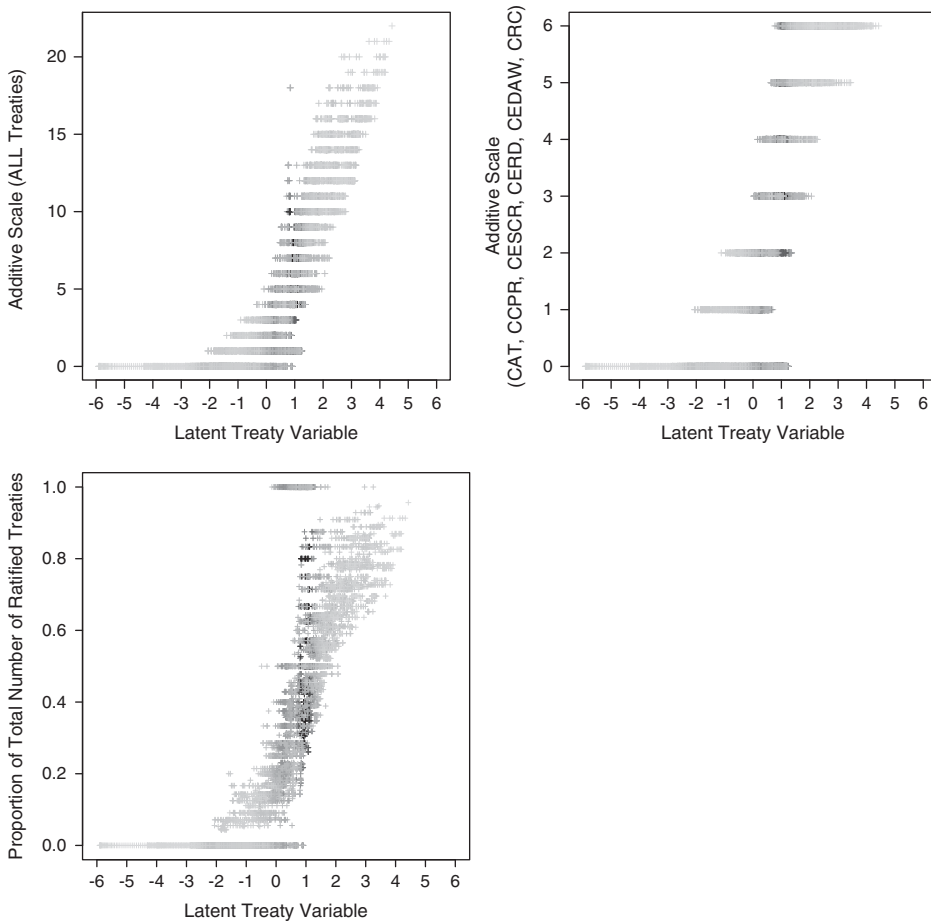


Fig. 1. Relationship between the latent treaty variable, two additive scales, and proportion

Notes: The upper left panel uses an additive scale that ranges from 0 to 23 and is based on the total number of ratified treaties and optional protocols contained in Table 1. The upper right panel uses an alternative additive scale that ranges from 0 to 6 and is based on a total of up to six ratified treaties. The selected treaties are commonly studied in the international relations literature and include the Convention Against Torture (CAT), the Covenant on Civil and Political Rights (CCPR), the Covenant on Economic, Social and Cultural Rights (CESCR), the Convention on the Elimination of All Forms of Racial Discrimination (CERD), the Convention on the Elimination of all Forms of Discrimination against Women (CEDAW), and the Convention on the Rights of the Child (CRC). The correlation coefficients between the new latent variable and the additive scales are 0.795 [95% CI:0.792,0.799] and 0.754 [95% CI:0.750,0.757] respectively. The lower left panel shows the relationship between the latent variable and proportion of ratified treaties for those available each year. The correlation coefficients between these variables is 0.742 [95% CI:0.738,0.746]. We generated 95% credible intervals by taking 1,000 draws from the posterior distribution of latent treaty variables, which were then used to estimate the distribution of correlation coefficients.

latent variable, which is based on all of the treaty variables contained in Table 1 to two alternative additive scales and a proportion variable. All of these variables are closely related, though each make different assumptions.<sup>63</sup> However, as Figure 1 illustrates, there is considerable overlap along

<sup>63</sup> In the case of the additive scales, the item weights are equal in each year (missing values from earlier years are assumed to be 0). For the proportion variable, the item weights change from year to year as additional treaties

the latent variable for cases that are assumed to be different along the additive scales. The upper left panel compares the latent variable to an additive scale that ranges from 0 to 23 and is based on the total number of ratified treaties, optional protocols, and articles contained in Table 1. The upper right panel uses an alternative additive scale that ranges from 0 to 6 and is based on the total number of up to six ratified treaties. The selected treaties are commonly studied in the human rights literature and include the Convention Against Torture (CAT), the Covenant on Civil and Political Rights (CCPR), the Covenant on Economic, Social and Cultural Rights (CESCR), the Convention on the Elimination of All Forms of Racial Discrimination (CERD), the Convention on the Elimination of all Forms of Discrimination against Women (CEDAW), and the Convention on the Rights of the Child (CRC). The lower left panel compares the latent variable with the proportion of ratified treaties of those that are available to ratify in a given year. To generate the additive scales and the proportion variable, I recode missing values as 0 before summing across the indicators, which is the common approach when using the additive scale in the treaty compliance literature.<sup>64</sup> Note that this is not a necessary choice in the context of a latent variable model because this model can account for missing data. Again, the missing data have substantive meaning, which unfortunately is ignored by scholars using only an additive scale.

Not surprisingly, embeddedness within the international human rights regime has increased since the end of the Second World War. Section A of the supplementary appendix displays numerically and visually the model parameters that link the latent treaty variable  $\theta$  to the observed binary treaty variables. Figure 2 displays the average level of embeddedness over time (see Section B in the supplementary appendix for more figures). In the next section I compare these treaty variables and the six binary treaty variables mentioned above with the two latent human rights variables.<sup>65</sup>

### *Measuring the Level of Respect for Human Rights*

Here, I briefly describe the two sets of latent human rights estimates, which were generated with models similar to the latent treaty variable defined above.<sup>66</sup> These models are dynamic in the estimate of the latent variable, just like the latent treaty model presented above, but include ordered logit functions that link the ordered categorical data derived from human rights reports with the latent human rights estimates.<sup>67</sup> In the next section, I compare these two latent human rights variables with ten treaty ratification variables.

To reiterate, latent variable models, with their focus on the theoretical relationship between data and model parameters, offer a principled way to bring together different pieces of information, even if that information is biased. The information in the *The Country Reports on*

(Footnote continued)

and protocols enter the dataset. Thus, the item weight change from year to year which makes over time comparisons problematic because it is not clear how to compare two countries of the same value on the proportion measure as the total number of treaties changes over time. A '1' on the proportion variable mean something very different when only one or two treaties are available for signature compared to years when all twenty-three indicators are open for ratification. The latent variable model used in this article provides a principled and open way to generate these item weights and then evaluate them.

<sup>64</sup> Hafner-Burton and Tsutsui 2005.

<sup>65</sup> Fariss 2014.

<sup>66</sup> These latent human rights estimates are derived from a latent variable model developed by Schnakenberg and Fariss (2014) and an extended version of this model that incorporates the standard of accountability by Fariss (2014).

<sup>67</sup> Again latent variable models are as simple as linear regression. The latent variable model I use in this paper to link binary treaty data together is as simple as logistic regression. The latent human rights variable model is as simple as ordered-logistic regression.

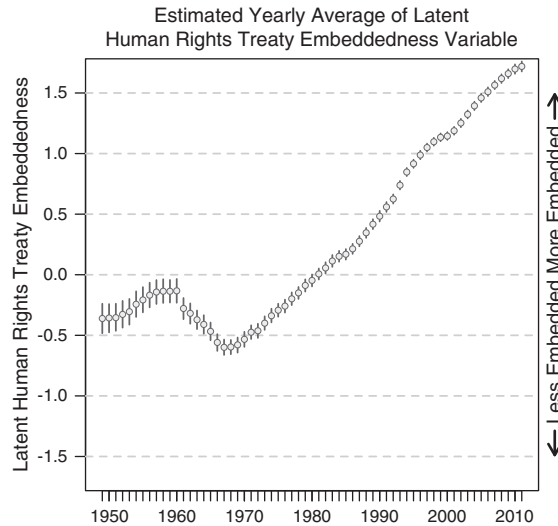


Fig. 2. Average level of the human rights treaty variable over time

Notes: On average, countries become increasingly embedded in the global human rights regime over time.

*Human Rights Practices* published annually by the US State Department and *The State of the World's Human Rights* report published annually by Amnesty International is reflective of both the historical context in which these reports were written and the true level of abuse in a given country during that period. The latent variable model I developed in 2014 provides a principled and transparent (though computationally complex) model that can account for these changes over time, which removes the temporal bias from the resulting latent variable estimates.<sup>68</sup>

The manifestation of the changing standard of accountability has been recognized by a substantial body of human rights scholars who have developed in-depth knowledge of specific cases.<sup>69</sup> As Sikkink notes, monitoring agencies and others 'have expanded their focus over time from a narrow concentration on direct government responsibility or the death, disappearance, and imprisonment of political opponents to a wider range of rights, including the right of people to be free from police brutality and the excessive use of lethal force'.<sup>70</sup> As I have argued, monitoring agencies are looking harder for abuse, they are looking in more places, and they are classifying more acts as abuse.<sup>71</sup> This argument is corroborated by several authors who suggest that the focus on police brutality, and the specific departments and locations of the abuse, in places like Brazil and Argentina are a relatively new aspect of human rights monitoring.<sup>72</sup> Marchesi and Sikkink specifically discuss the pattern of observation and the increasing monitoring that occurred in Brazil as it transitioned from a military regime into a democracy during the 1970s and 1980s:

This focus on police killings and torture has come to constitute the main focus of reports on Brazil throughout the 1990s and 2000s. The question is not whether this is a very serious violation of human rights. The question from the point of view of a researcher sensitive to information effects is

<sup>68</sup> Fariss 2014.

<sup>69</sup> Brysk 1994; Clark and Sikkink 2013; Keck and Sikkink 1998; Marchesi and Sikkink 2015; Sikkink 2011.

<sup>70</sup> Sikkink (2011, p. 159).

<sup>71</sup> Fariss 2014.

<sup>72</sup> See Fariss 2014; Marchesi and Sikkink 2015; Sikkink 2011.

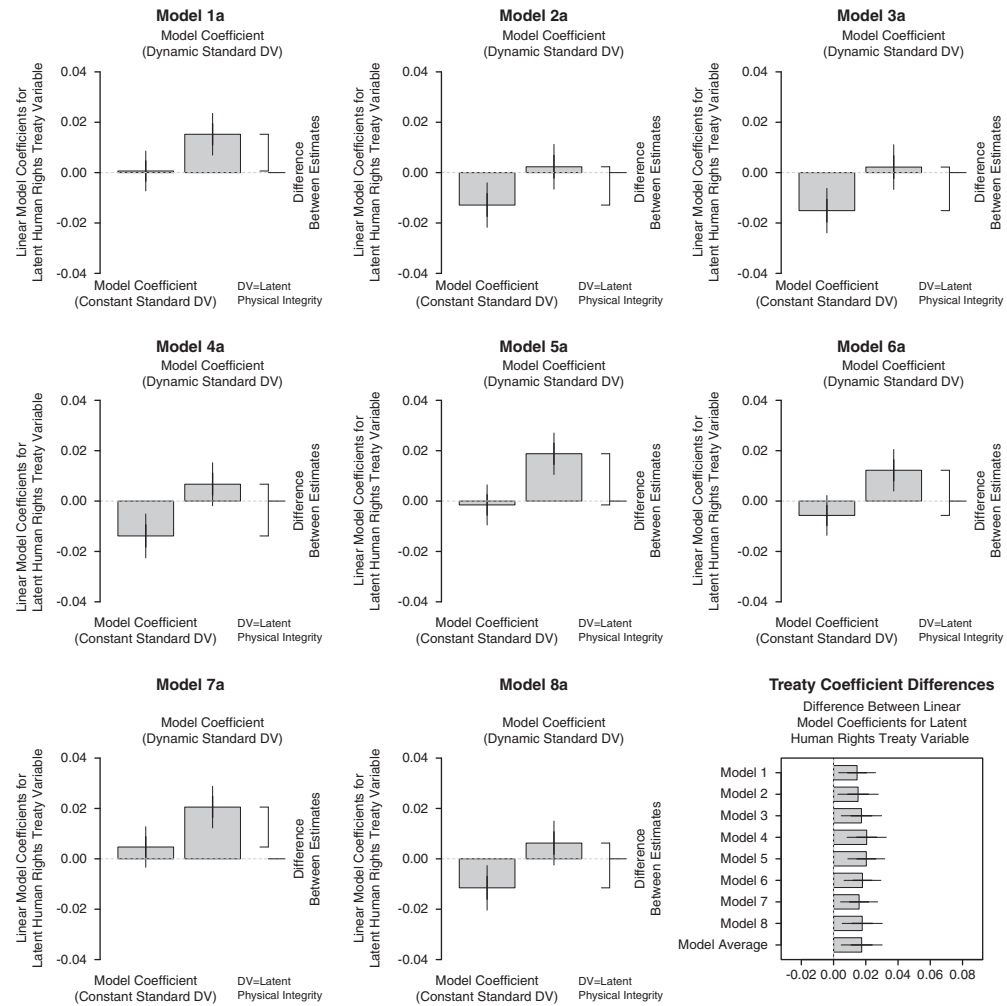


Fig. 3. Models for all human rights treaties using human rights latent treaty variable

Notes for Fig. 3: Estimated coefficient from the linear models using the dependent latent physical integrity variables from the constant standard model and the dynamic standard model respectively. The thick lines represent  $1 \pm$  the standard error of the coefficient. The thin lines represent  $2 \pm$  the standard error of the coefficient. The differences are all similar across models and all statistically different from 0. See the tables in Section F of the supplementary appendix for full model results.

‘Do the Brazilian police kill and mistreat more victims today than they did in the 1970s and 1980s or do we know more about that killing and mistreatment today than we did before?’ We think it is quite possible that in the 1970s and 1980s, when the Brazilian government had large numbers of political prisoners and was killing and disappearing political opponents, the Brazilian police were also very violent with regard to criminal suspects, but such violence was not documented or reported.<sup>73</sup>

Overall, there is a substantial amount of qualitative evidence that suggests that the standard of accountability has changed over time. The latent variable model I have developed formalizes

<sup>73</sup> Marchesi and Sikkink (2015, pp. 21–2).

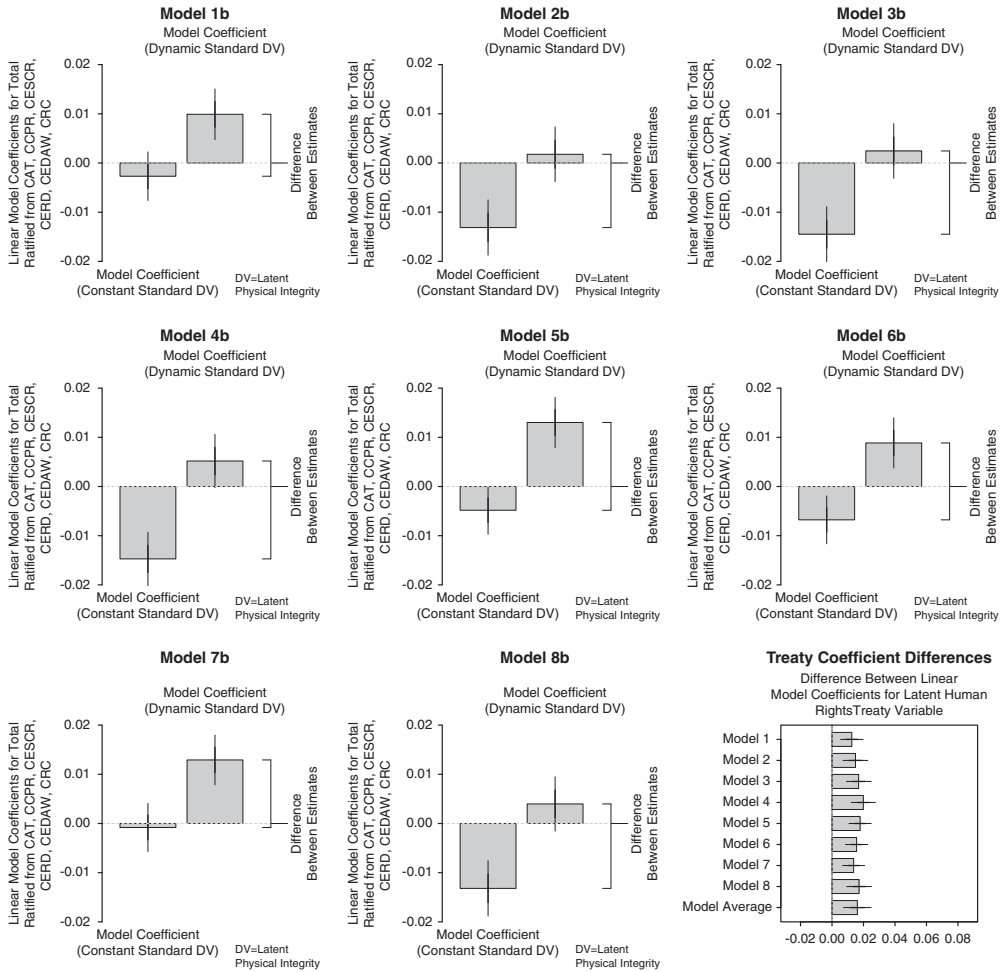


Fig. 4. Models for selected human rights treaties (CAT, CCPR, CESC, CERD, CEDAW, CRC), using human rights count treaty variable  
See Notes for Fig. 3.

this concept and generates new unbiased estimates of the level of respect for human rights by taking it into account.<sup>74</sup> The latent variable model incorporates information from thirteen original data sources. The thirteen variable names, temporal coverage, and data type of each variable are displayed in Section C of the supplementary appendix.<sup>75</sup>

It is important to note that the version of the latent variable that does not account for the changing standard of accountability (constant standard model) is consistent with the

<sup>74</sup> Fariss 2014.

<sup>75</sup> Names, operationalizations, citations, and data sources for each of the thirteen original variables are displayed in Section C of the supplementary appendix. The other measures included in the latent human rights variable estimates include an additional ordered variable of torture (Conrad, Haglund and Moore 2013), a binary measure of government one-sided killings adapted from Eck and Hultman (2007), three measures of genocide/politicide (Harff 2003; Harff and Gurr 1988; Rummel 1994; Rummel 1995), and a binary measure of political executions adapted from Taylor and Jodice (1983).

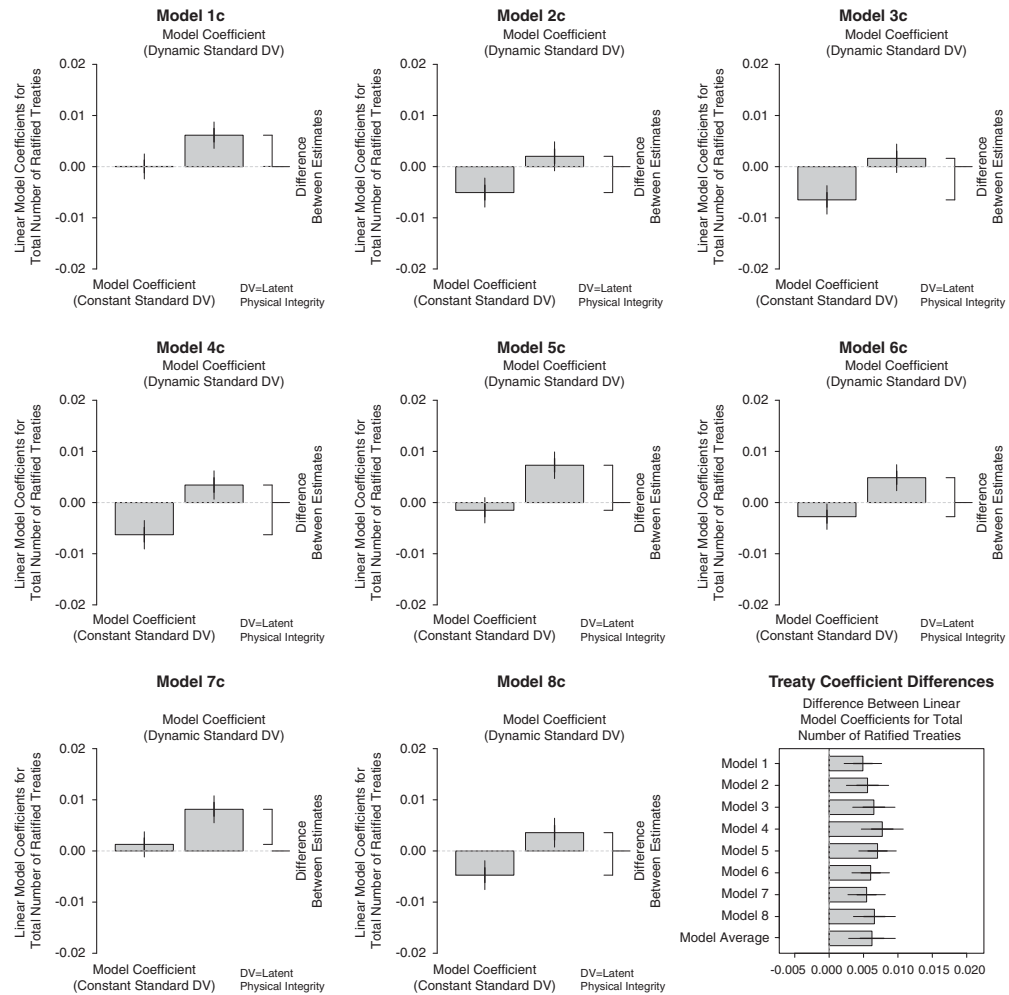


Fig. 5. Models for all human rights treaties, using human rights proportion treaty variable  
See Notes for Fig. 3.

assumptions used by all existing human rights scales.<sup>76</sup> For the latent human rights variable that does incorporate the changing standard of accountability (dynamic standard model), the event-based variables act as a consistent baseline with which to compare the levels of the standards-based variables. The event-based variables act as this baseline because they are each updated as new information about the specific events becomes available and are based on evidence from several regularly supplemented primary source documents (see Section C in the supplementary appendix for a list of sources); whereas, the human rights reports are produced in a specific historical context and never updated. Moreover, the producers of these events-based variables are focused on the extreme end of the repression spectrum (e.g., genocide, politicide, mass-repression, one-sided government killings). This focus makes identifying these events a relatively easier task than the more difficult one of observing behaviors such as torture.

<sup>76</sup> Cingranelli and Richards 1999; Cingranelli, Richards and Clay 2015a; Cingranelli, Richards and Clay 2015b; Gibney, Cornett and Wood 2012; Hathaway 2002.

TABLE 2 *Summary of Visual Displays of Regression Results for Nine Treaty Variables*

	Human Rights Treaty Variable Description
Figure 3	Latent Human Rights Treaty Variable
Figure 4	Count of Selected Human Rights Treaties (CAT, CCPR, CESC, CERD, CEDAW, CRC)
Figure 5	Count of All Human Rights Treaties (see Table 1)
Figure 6	Proportion of All Human Rights Treaties Available for Ratification in Year $t$ (see Table 1)
Figure 7	Convention Against Torture (CAT)
Figure 8	Convention on the Elimination of all Forms of Discrimination Against Women (CEDAW)
Figure 9	International Covenant on Civil and Political Rights (CCPR)
Figure 10	International Covenant on Economic, Social, and Cultural Rights (CESC)
Figure 11	International Convention on the Rights of the Child (CRC)
Figure 12	International Convention on the Elimination of All Forms of Racial Discrimination (CERD)

*Note:* Summary information for Figures 7–12 is contained in Table 2. Each figure displays linear regression coefficients for one of two dependent variables regressed on the selected treaty variable and controls. Each treaty variable is included in each of eight model specifications described above. The difference between the treaty coefficients is similar across model specifications for all treaty variables. To reiterate the results: even though the individual coefficients change depending on the model specification, the differences are consistent, which is a substantively important finding that eliminates concern that the use of a particular control variable is driving the results. And again, the results always contradict the negative findings from existing research. The coefficient for the each of the various treaty variables flip signs in every model permutation presented across the figures.

In areas in which large scale repressive events occur, lower level abuses are often missed or simply ignored.<sup>77</sup> As Brysk notes, ‘Incidents of kidnapping and torture which would register as human violations elsewhere did not count in Argentina. The volume of worse rights abuses set a perverse benchmark and absorbed monitoring capabilities’.<sup>78</sup> These features suggest that the event-based data are a valid representation of the historical record to date, which again, acts as a consistent baseline by which to compare how the changing standard of accountability is associated with the standard-based variables over time. The changing standard of accountability is incorporated into the latent variable model of human rights by allowing the baseline probability of a human rights variable to be coded at a specific value to change over time.<sup>79</sup>

The supplementary appendix (Section D) includes a graph that shows the difference in the two latent human rights variables each year, which illustrates the new picture that emerges of improving physical integrity practices over time. The supplementary appendix (Section E) also includes a ten-country time series graph comparing the two latent human rights estimates.

#### RESULTS: THE RELATIONSHIP BETWEEN HUMAN RIGHTS AND TREATY RATIFICATION

In this section, I illustrate the substantive importance of the changing standard of accountability for understanding human rights compliance over time by showing that ratification of UN human rights treaties and respect for physical integrity rights is positive. The results contradict negative findings

<sup>77</sup> Brysk 1994.

<sup>78</sup> Brysk (1994, p. 681).

<sup>79</sup> See Fariss (2014) for more details on the formalization of this part of the latent variable model.

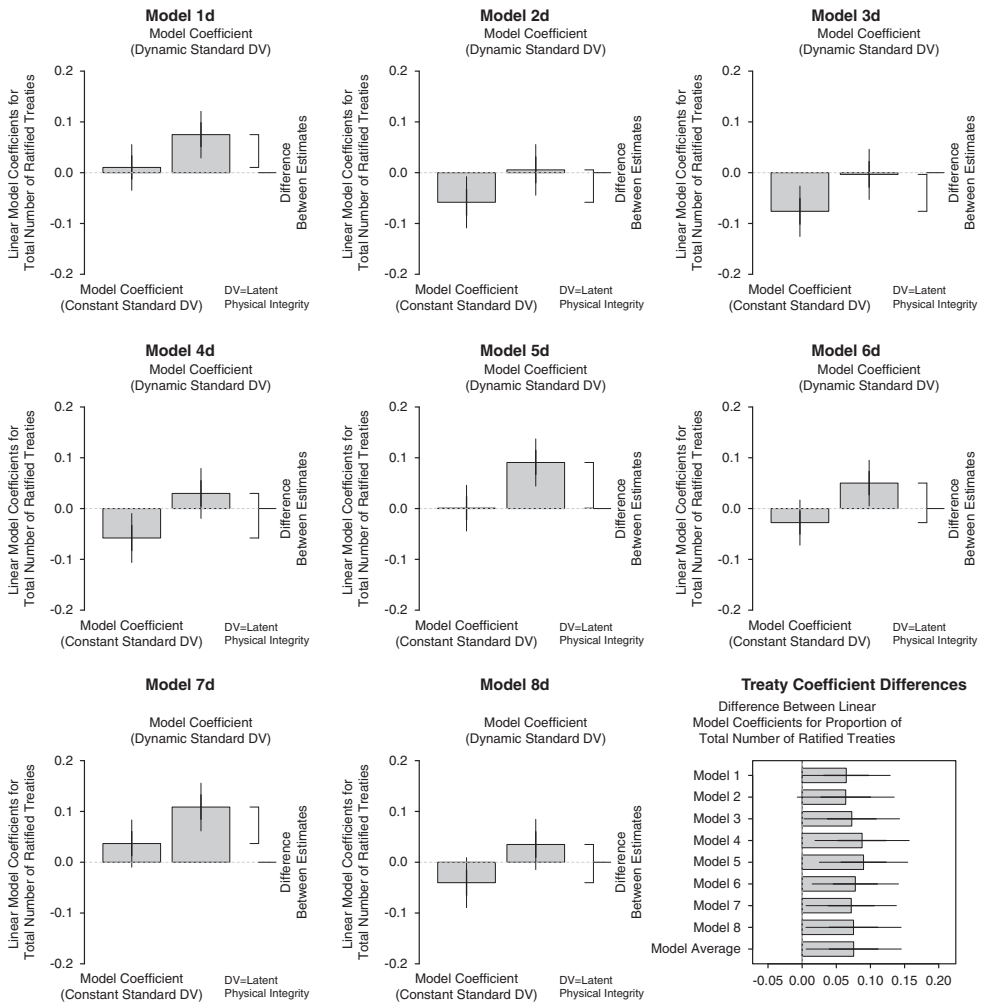


Fig. 6. Proportion of all human rights treaties available for ratification in year  $t$ , using human rights count treaty variable  
See Notes for Fig. 3.

from existing research. As the standard of accountability increases over time, empirical associations with human rights data derived from standards-based documents and other variables will become increasingly biased if changes in the human rights documents are not considered. This is especially true for variables that measure the existence of institutions that are correlated with time such as whether or not a particular treaty like the UN Convention Against Torture has been ratified or not.

To simplify the presentation of results, I visually compare two linear model coefficients using the dependent variable from the latent variable model that does not account for the changing standard of accountability (labeled the constant standard model) and the dependent variable from the latent variable model that does account for the changing standard of accountability (labeled the dynamic standard model). I estimate two linear regression equations using the latent physical integrity variables from the two measurement models. I run a regression of these variables on several treaty variables, including the latent treaty variable defined above, two

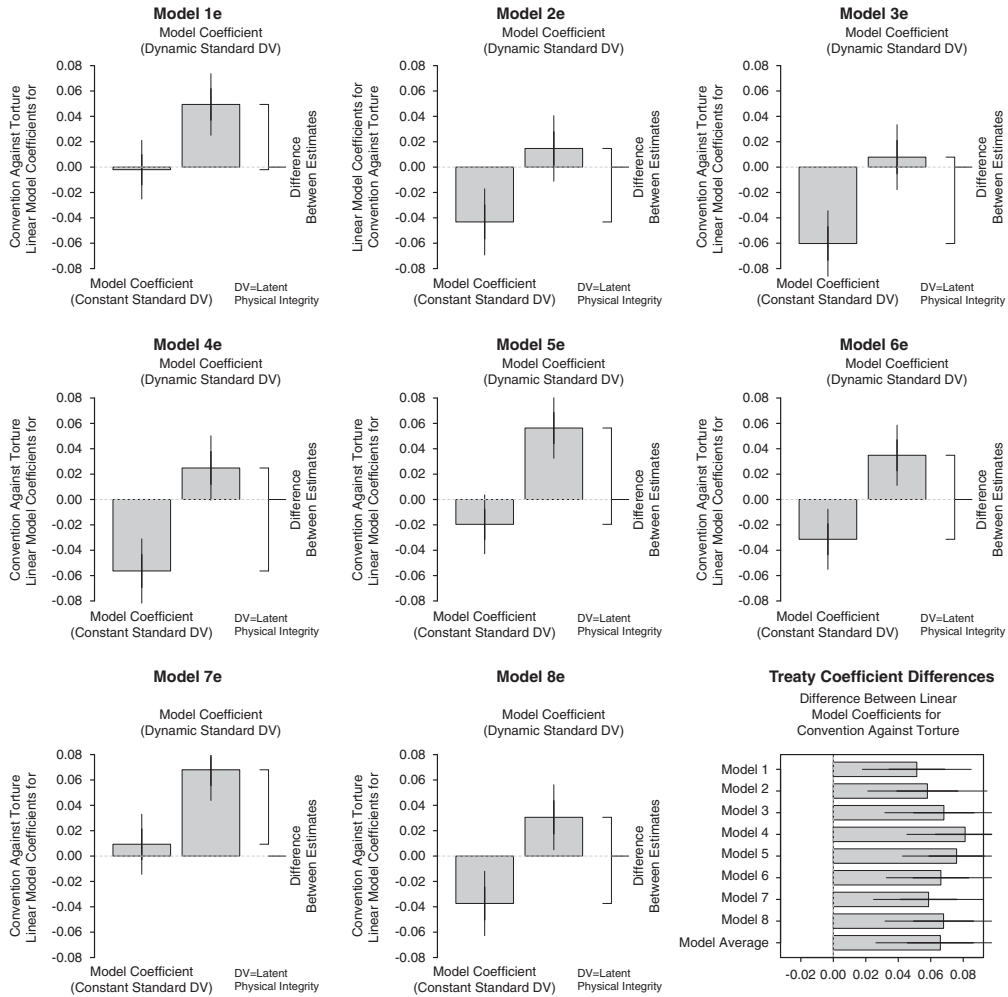


Fig. 7. Models for the Convention Against Torture using binary treaty variable  
See Notes for Fig. 3.

versions of an additive treaty scale, and six binary variables. Each binary treaty variable measures whether or not a country has ratified the Convention Against Torture (CAT), the Convention on the Elimination of All Forms of Discrimination Against Women (CEDW), the Covenant on Civil and Political Rights (CCPR), Covenant on Civil and Political Rights (CCPR), Convention on the Rights of the Child (CRC), or Convention on the Elimination of All Forms of Racial Discrimination (CERD) in a given year. I also include several control variables in eight different specifications.<sup>80</sup> Each model always includes the lagged version of one of the two human rights variables and the lagged version of one of the treaty variables.

<sup>80</sup> Control variables include a measure of democracy (Marshall, Jaggers and Gurr 2013), the natural log of GDP per capita (Gleditsch 2002), the natural log of population (Gleditsch 2002), and the lagged value of the latent human rights variable and finally the lagged value of one of the various different treaty variables. Overall, the choice of variables for these models does not change the difference in the relationship between treaty ratification and respect for human rights.

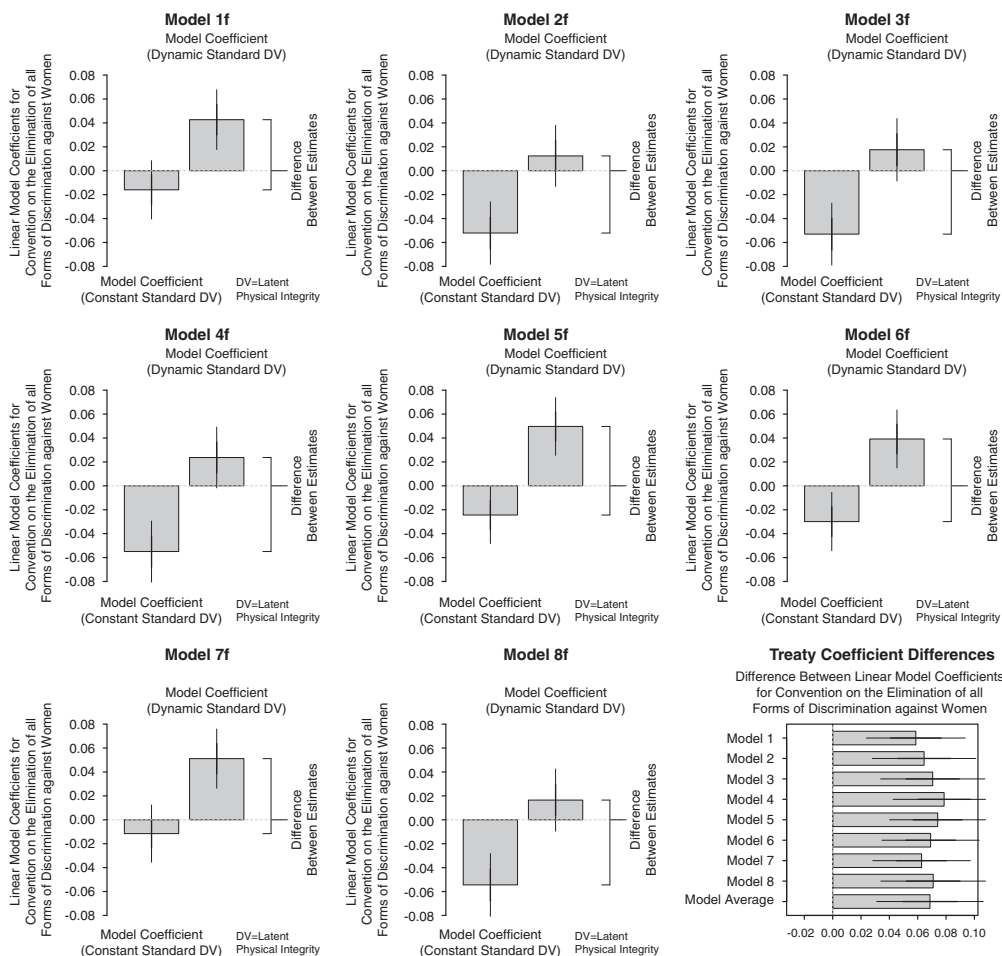


Fig. 8. Models of the Convention on the Elimination of All Forms of Discrimination Against Women, using binary treaty variable  
See Notes for Fig. 3.

The model comparisons demonstrate that the differences between the coefficients are similar across all of the model specifications; adding or removing any specific control variable does not change the difference between the coefficients. Thus, the results always contradict the negative findings from existing research.<sup>81</sup> That is, omitted variable bias does not change the substantive meaning of the difference in the relationship between treaty ratification and respect for human rights. The eight linear regression models are specified as follows:

$$\text{Model 1 } y_{it} \sim \beta_0 + \beta_1 * y_{i,t-1} + \beta_2 * \text{treaty}_{t-1}$$

$$\text{Model 2 } y_{it} \sim \beta_0 + \beta_1 * y_{i,t-1} + \beta_2 * \text{treaty}_{t-1} + \beta_3 * \text{Polity2}_{t-1}$$

$$\text{Model 3 } y_{it} \sim \beta_0 + \beta_1 * y_{i,t-1} + \beta_2 * \text{treaty}_{t-1} + \beta_3 * \text{Polity2}_{t-1} + \beta_4 * \ln(\text{gdppc}_{t-1})$$

<sup>81</sup> The difference between the treaty variable coefficients from any two competing modes is based on the following Z-score:  $\frac{\beta_{\text{dynamic}} - \beta_{\text{constant}}}{\sqrt{SE(\beta_{\text{dynamic}})^2 + SE(\beta_{\text{constant}})^2}}$ .

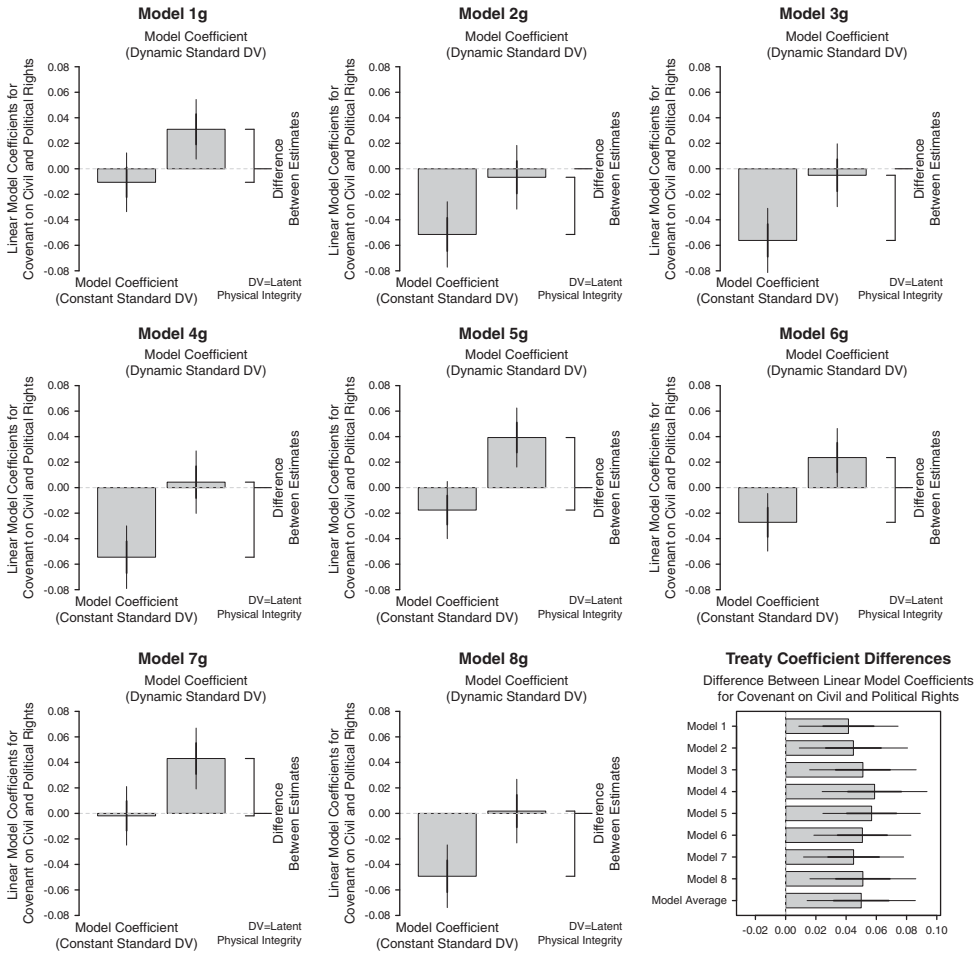


Fig. 9. Models of the International Covenant on Civil and Political Rights, using binary treaty variable  
See Notes for Fig. 3.

$$\text{Model 4 } y_{it} \sim \beta_0 + \beta_1 * y_{i,t-1} + \beta_2 * \text{treaty}_{t-1} + \beta_3 * \text{Polity2}_{t-1} + \beta_4 * \ln(\text{gdppc}_{t-1}) + \beta_5 * \ln(\text{population}_{t-1})$$

$$\text{Model 5 } y_{it} \sim \beta_0 + \beta_1 * y_{i,t-1} + \beta_2 * \text{treaty}_{t-1} + \beta_4 * \ln(\text{gdppc}_{t-1}) + \beta_5 * \ln(\text{population}_{t-1})$$

$$\text{Model 6 } y_{it} \sim \beta_0 + \beta_1 * y_{i,t-1} + \beta_2 * \text{treaty}_{t-1} + \beta_4 * \ln(\text{gdppc}_{t-1})$$

$$\text{Model 7 } y_{it} \sim \beta_0 + \beta_1 * y_{i,t-1} + \beta_2 * \text{treaty}_{t-1} + \beta_5 * \ln(\text{population}_{t-1})$$

$$\text{Model 8 } y_{it} \sim \beta_0 + \beta_1 * y_{i,t-1} + \beta_2 * \text{treaty}_{t-1} + \beta_3 * \text{Polity2}_{t-1} + \beta_5 * \ln(\text{population}_{t-1})$$

$y_{it}$  is the country–year human rights variable generated from either the dynamic standard model or the constant standard model described above. Each regression model is estimated 2 by 3 by 8 times in order to compare the two competing dependent variables, using each of the three different treaty variables (the latent treaty variable and the two scales) within eight different model specifications. I also consider six additional binary treaty variables. All models include the same set of country–year observations from 1965, the year the International Convention on the Elimination of All Forms of Racial Discrimination opened for signature, through 2010. The results are also consistent for the period 1949 through 2010 and 1976 through 2010.

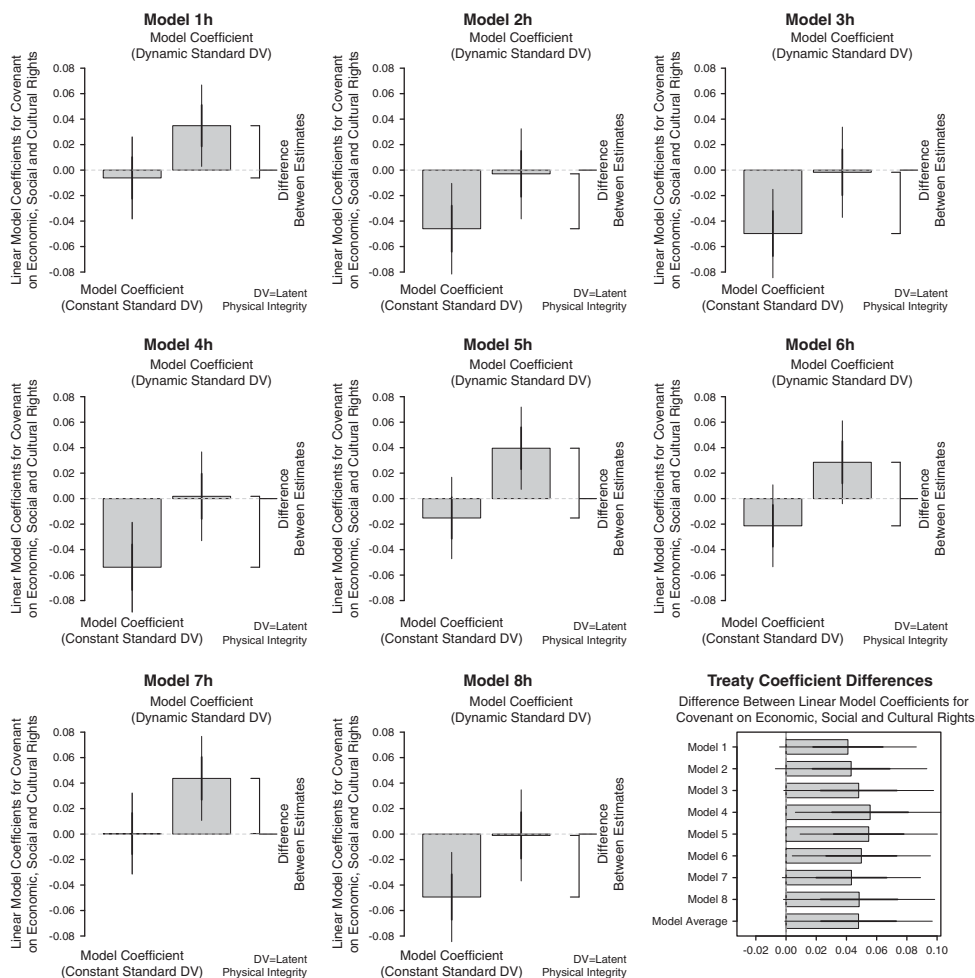


Fig. 10. Models of the International Covenant on Economic, Social, and Cultural Rights, using binary treaty variable

See Notes for Fig. 3.

The key piece of information to consider from the various models is the difference between the coefficients estimated for the treaty variable for the two competing dependent variables. Each of the figures below visually display the two coefficients estimates for all eight model specifications using one of the treaty variables. The ninth panel in each figure displays the difference between these coefficients across the eight model specifications. This panel in each set of figures demonstrates that the differences between regression coefficients are statistically the same no matter what additional variables enter the model specification.

Though the individual coefficients for the treaty variables change across the eight models, the differences between the coefficients are consistent across all eight specifications for all treaty variables. By estimating all of the various model specifications for both dependent variables, I am able to demonstrate that the estimated difference between the coefficients using the corrected human rights variable (dynamic standard model) compared to the uncorrected human

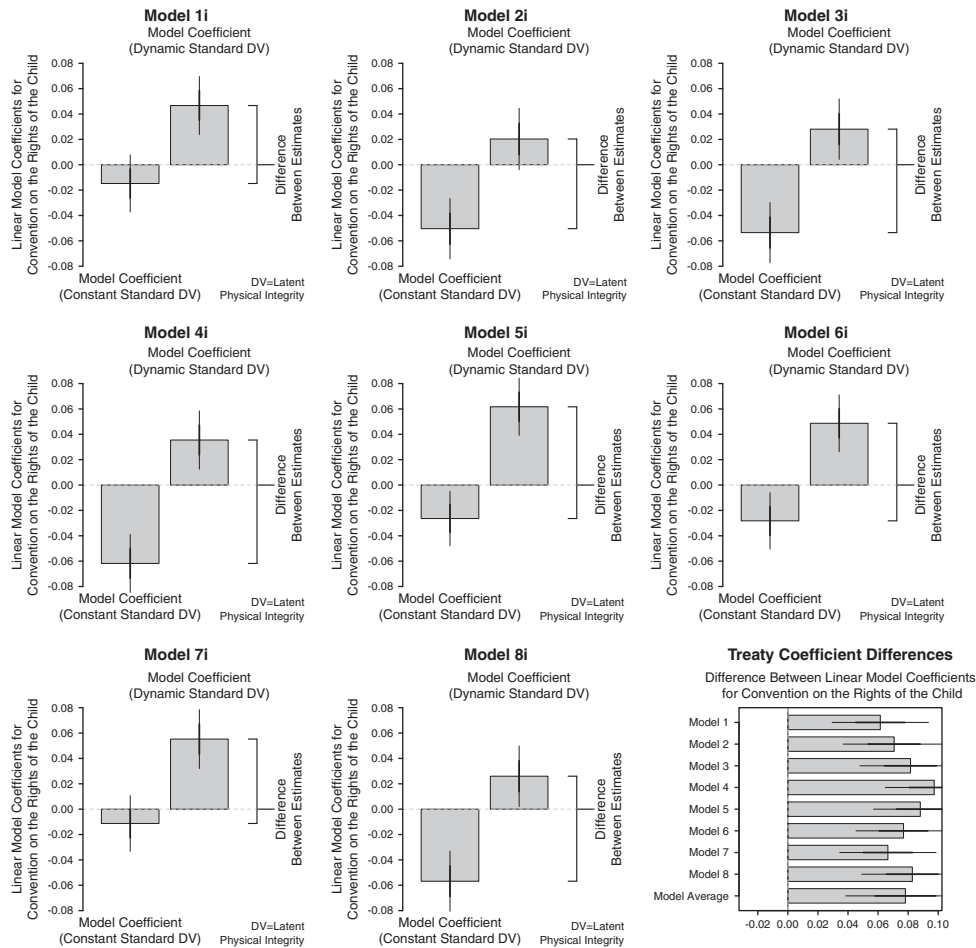


Fig. 11. Models of the International Convention on the Rights of the Child, using binary treaty variable See Notes for Fig. 3.

rights variable (constant standard model) are consistent and statistically different from 0. Thus, even though the individual coefficients change depending on the model specification, the differences are consistent, which is a substantively important finding that eliminates concern that the use of a particular control variable is driving the results. The differences between coefficients are therefore robust to variable selection. The coefficient for the various treaty variables flip signs in every model permutation presented across the figures and therefore contradict the negative findings from existing research. These results again suggest that human rights treaties are not simply cover for human rights abusers.

Overall, new relationships between treaty ratification and the level of human rights respect are obtained by replacing the dependent variable derived from the constant standard model with the one from the dynamic standard model, no matter what the specification of the regression model. Recall that the dynamic standard model accounts for the changing standard of accountability, while the constant standard model does not. For example, Figure 3 plots the linear model coefficient for the latent rights treaty variable. Again, each model uses one of the

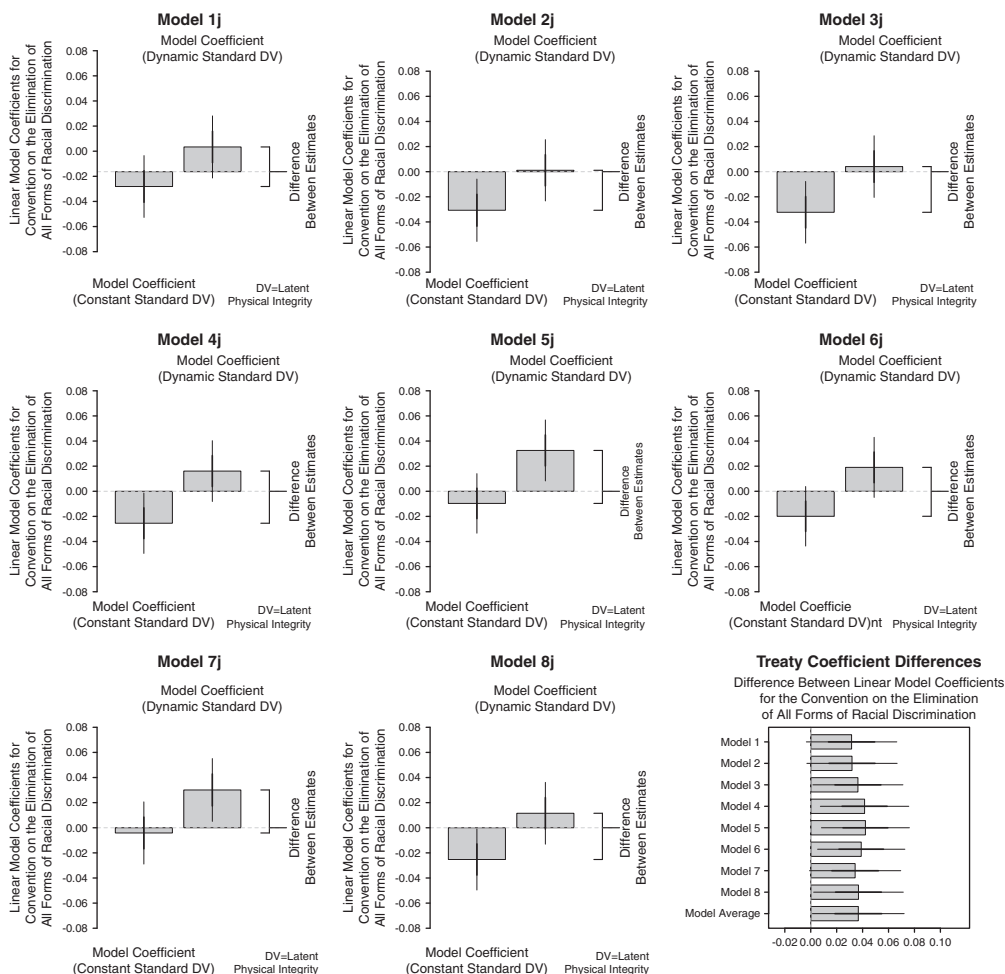


Fig. 12. Models of the International Convention on the Elimination of All Forms of Racial Discrimination, using binary treaty variable  
See Notes for Fig. 3.

two latent physical integrity dependent variables and various control variables. Figures 4 and 5 display results using the two additive treaty scales. Figure 6 displays results using the treaty proportion variable. I also present visual results for the coefficients for the binary variable measures for the CAT, CEDAW, CCPR, CESC, CRC, and CERD ratification variables respectively.<sup>82</sup>

<sup>82</sup> The models presented in this article are not designed for causal inference, though a variety of selection issues are known to exist. Recent attempts to address the selection issue include Neumayer (2005), Simmons and Hopkins (2005), Von Stein (2005), Simmons (2009), Hill (2010), and most recently Lupu (2013b). It is important to note though, that the selection issue is unrelated to the differences in the two latent human rights variables used in the regressions results presented here. Moreover, efforts to address the selection issue were driven in large part by the counterintuitive negative findings, which are an artifact of the unaccounted for changing standard of accountability, not the actual levels of respect for human rights.

TABLE 3 *Comparison of Results with Other Studies\**

Treaty ratification variable	Positive relationship	Negative relationship
Latent Human Rights Treaty Variable	This article (Figure 3)	–
Count of Selected Human Rights Treaties (CAT, CCPR, CESC, CERD, CEDAW, CRC)	This article (Figure 4)	Hafner-Burton and Tsutsui (2005)
Count of All Human Rights Treaties	This article (Figure 5)	–
Proportion of All Human Rights Treaties Available for Ratification in Year <i>t</i>	This article (Figure 6)	–
Convention Against Torture (CAT)	This article (Figure 7), Cole (2012), Conrad and Ritter (2013), Fariss (2014), Hill and Jones (2014), <sup>†</sup> Landman (2005), Lupu (2015)	Hafner-Burton and Tsutsui (2005), Hafner-Burton and Tsutsui (2007), Hathaway (2002), Hill and Jones (2014), <sup>†</sup> Hollyer and Rosendorff (2011), Neumayer (2005)
Convention on the Elimination of All Forms of Discrimination Against Women (CEDAW)	This article (Figure 8), Hill (2010), Lupu (2013a), Simmons (2009)	Hafner-Burton and Tsutsui (2005)
Covenant on Civil and Political Rights (CCPR)	This article (Figure 9), Landman (2005), Lupu (2013b), Hill and Jones (2014), <sup>†</sup> Simmons (2009)	Hafner-Burton and Tsutsui (2005), Hafner-Burton and Tsutsui (2007), Hill (2010), Hill and Jones (2014) <sup>†</sup>
Covenant on Economic, Social and Cultural Rights (CESCR)	This article (Figure 10)	Hafner-Burton and Tsutsui (2005)
Convention on the Rights of the Child (CRC)	This article (Figure 11)	Hafner-Burton and Tsutsui (2005)
Convention on the Elimination of All Forms of Racial Discrimination (CERD)	This article (Figure 12)	Hafner-Burton and Tsutsui (2005)

\*The findings presented in this study corroborate and strengthen the positive results present in many other recent studies that think carefully about research design strategies. These studies are presented in this table and discussed throughout the text of this article. Overall, the findings in these studies are strengthened because of the measurement issues addressed in this article. The negative results presented in other studies are likely only attributable to the systematic measurement bias caused by the changing standard of accountability. Several other studies report neutral results when exploring the relationship between the ratification of different treaties and respect for human rights (e.g. Keith 1999). Again, the negative and neutral findings are most likely due to the unaccounted for changing standard of accountability implicitly contained in the data used in these earlier studies.

<sup>†</sup>Hill and Jones (2014) find both positive and negative correlations between treaty ratification and respect for human rights, which is why this study is listed in both columns of this table. Hill and Jones (2014) find positive relationships using the latent variable estimates that account for the changing standard of accountability developed by Fariss (2014) but find negative relationships using the standard additive human rights scales (i.e., Cingranelli and Richards 1999; Cingranelli, Richards and Clay 2015a; Cingranelli, Richards and Clay 2015b; Gibney, Cornett and Wood 2012).

## CONCLUSION

The results presented in this article represent a first step towards re-evaluating what has become conventional wisdom in the literature of international treaty compliance and human rights more generally.<sup>83</sup> I have presented evidence that the ratification of human rights treaties is empirically associated with higher levels of respect for human rights over time and across countries. This evidence bolsters claims that the negative correlations between variables measuring respect for human rights and variables measuring treaty ratification are an artifact of some other unaccounted for process.<sup>84</sup> With human rights data that now accounts for the changing standard of accountability, a new picture has emerged of improving levels of respect for human rights, which coincides with the increasing embeddedness of countries within the international human rights regime. This positive relationship is robust to a variety of measurement strategies and model specifications. These results are important because much of the theorizing in international relations begins with the premise that international human rights treaties are not effective in order to explain these puzzling, negative correlations<sup>85</sup> or to provide policy recommendations based on these anomalous findings.<sup>86</sup> The results presented in this article, summarized in Table 3, cast considerable doubt on the empirical bases of these other research projects. Thus, conclusions and policy recommendations from this earlier research need to be re-evaluated.<sup>87</sup>

Recent work by some human rights scholars has begun to unpack the institutional mechanisms by which international human rights treaty commitments can become effective.<sup>88</sup> With innovative new research designs available<sup>89</sup> and now new data and measurement tools as well,<sup>90</sup> scholars have the ability to begin the systematic reassessment of the role that international human rights treaties play in mitigating the use of repressive tactics by all types of governments, working under a variety of institutional designs.

In closing, I wish to emphasize that a science of human rights requires valid comparisons of repression levels across different spatial and temporal contexts.<sup>91</sup> Scholars of human rights,

<sup>83</sup> In a recent article, Hill and Jones (2014) use cross-validation and random forest methods to determine the predictive power of the covariates identified as important in the literature on human rights using the existing CIRI and PTS physical integrity scales and the new latent human rights variable that adjusts for the changing standard of accountability. The cross-validation and random forest methods corroborate the results that ratification of the Convention Against Torture and the International Covenant on Civil and Political Rights are positively associated with respect for human rights. These authors do not consider any other UN human rights treaty variables, however they do find that measures of civil war (Davenport 2007; Poe and Tate 1994; Poe, Tate and Keith 1999), 'youth bulges' (Nordas and Davenport 2013), domestic legal institutions (Crabtree and Fariss 2015; Keith 2012; Keith and Poe 2004; Keith, Tate and Poe 2009), and state reliance on natural resource rents (Demeritt and Young 2013), are good predictors on levels of repression.

<sup>84</sup> Clark and Sikkink 2013; Fariss 2014; Goodman and Jinks 2003; Marchesi and Sikkink 2015.

<sup>85</sup> Hafner-Burton 2013; Hafner-Burton, Tsutsui and Meyer 2008; Hollyer and Rosendorff 2011; Posner 2014; Vreeland 2008.

<sup>86</sup> Hopgood 2013; Moyn 2010; Posner 2014.

<sup>87</sup> See, for example, Dancy and Fariss (2015).

<sup>88</sup> Simmons (2009) looks specifically at the enforcement power of domestic institutions, whereas Conrad and Ritter (2013), Cole and Ramirez (2013), Lupu (2013a), Powell and Staton (2009), and Sikkink and her colleagues look at the intersection of domestic legal institutions and the specific human rights provisions within the various instruments that make up the international human rights treaty regime (e.g., Dancy and Sikkink 2012; Kim and Sikkink 2010; Sikkink 2011). Sandholtz (2012) builds on earlier work in the human rights literature by linking international treaties, and constitutional design, with respect for human rights (e.g., Davenport 1996; Keith 2002; Keith 2012; Keith and Poe 2004; Keith, Tate and Poe 2009; Mayerfeld 2016).

<sup>89</sup> Hill 2010; Hill and Jones 2014; Lupu 2013a; Lupu 2013b.

<sup>90</sup> Fariss 2014; Fariss and Schnakenberg 2014; Schnakenberg and Fariss 2014.

<sup>91</sup> Schnakenberg and Fariss 2014.

repression, and contentious politics are aware of the fact that the quality of information across different political contexts is messy. Heterogeneity in information quality and the availability of information are major stumbling blocks for generating valid inferences about the topics our research community cares about. Latent variable models, with their focus on the theoretical relationship between data and model parameters, offer a principled way to bring together this information and make sense of it. This approach to measurement is useful because it shifts the burden of validity from the primary source documentation and raw data to the model parameters that bind these diverse pieces of information together.

## REFERENCES

- Adcock, Robert, and David Collier. 2001. Measurement Validity: A Shared Standard for Qualitative and Quantitative Research. *American Political Science Review* 95 (3):529–46.
- Albert, James H., and Val E. Johnson. 1999. *Ordinal Data Modeling*. New York: Springer-Verlag.
- Barbera, Pablo. 2015. Birds of the Same Feather Tweet Together. Bayesian Ideal Point Estimation Using Twitter Data. *Political Analysis* 23 (1):76–91.
- Bond, Robert M., and Solomon Messing. 2015. Quantifying Social Media's Political Space: Estimating Ideology from Publicly Revealed Preferences on Facebook. *American Political Science Review* 109 (1):62–78.
- Bonica, Adam. 2012. Ideology and Interests in the Political Marketplace. *American Journal of Political Science* 57 (2):294–311.
- Borsboom, Denny. 2005. *Measuring the Mind*. Cambridge: Cambridge University Press.
- Brysk, Alison. 1994. The Politics of Measurement: The Contested Count of the Disappeared in Argentina. *Human Rights Quarterly* 16 (4):676–92.
- Caughey, Devin, and Christopher Warshaw. 2015. Dynamic Estimation of Latent Opinion Using a Hierarchical Group-Level IRT Model. *Political Analysis* 23 (2):197–211.
- Cingranelli, David L., and David L. Richards. 1999. Measuring the Level, Pattern, and Sequence of Government Respect for Physical Integrity Rights. *International Studies Quarterly* 43 (2):407–17.
- Cingranelli, David L., David L. Richards, and K. Chad Clay. 2015a. The Cingranelli-Richards (CIRI) Human Rights Data Project Coding Manual Version. Available from <http://www.humanrightsdata.com/p/data-documentation.html>, accessed 14 April 2014.
- . 2015b. The Cingranelli-Richards Human Rights Dataset Version. Available from <http://www.humanrightsdata.com/p/data-documentation.html>, accessed 14 April 2014.
- Clark, Ann Marie, and Kathryn Sikkink. 2013. Information Effects and Human Rights Data: Is the Good News about Increased Human Rights Information Bad News for Human Rights Measures? *Human Rights Quarterly* 35 (3):539–68.
- Clinton, Joshua, Simon Jackman, and Douglas Rivers. 2004. The Statistical Analysis of Roll Call Data. *American Political Science Review* 98 (2):355–70.
- Cole, Wade M. 2012. Human Rights as Myth and Ceremony? Reevaluating the Effectiveness of Human Rights Treaties, 1981–2007. *American Journal of Sociology* 117 (4):1131–71.
- Cole, Wade M., and Francisco O. Ramirez. 2013. Conditional Decoupling: Assessing the Impact of National Human Rights Institutions, 1981 to 2004. *American Sociological Review* 78 (4):702–25.
- Conrad, Courtenay R. 2014. Divergent Incentives for Dictators Domestic Institutions and (International Promises Not to) Torture. *Journal of Conflict Resolution* 58 (1):34–67.
- Conrad, Courtenay R., and Emily Hencken Ritter. 2013. Treaties, Tenure, and Torture: The Conflicting Domestic Effects of International Law. *Journal of Politics* 75 (2):397–409.
- Conrad, Courtenay R., Jillienne Haglund, and Will H. Moore. 2013. Disaggregating Torture Allegations: Introducing the Ill-Treatment and Torture (ITT) Country–Year Data. *International Studies Perspectives* 14 (2):199–220.
- Crabtree, Charles D., and Christopher J. Fariss. 2015. Uncovering Patterns among Latent Variables: Human Rights and De Facto Judicial Independence. *Research & Politics* 2 (3):1–9, doi:10.1177/2053168015605343

- Dancy, Geoff, and Christopher J. Fariss. 2015. Rescuing Human Rights Law from International Legalism and Its Critics; Tulane University and Penn State University working paper. <http://ssrn.com/abstract=2506144>.
- Dancy, Geoff, and Kathryn Sikkink. 2012. Ratification and Human Rights Prosecutions: Toward a Transnational Theory of Treaty Compliance. *NYU Journal of International Law and Politics* 44 (3):751–90.
- Davenport, Christian. 1996. Constitutional Promises and Repressive Reality: A Cross-National Time-Series Investigation of Why Political and Civil Liberties Are Suppressed. *Journal of Politics* 58 (3): 627–654.
- . 2007. State Repression and Political Order. *Annual Review of Political Science* 10:1–23.
- Demeritt, Jacqueline H.R., and Joseph K. Young. 2013. A Political Economy of Human Rights: Oil, Natural Gas, and State Incentives to Repress. *Conflict Management and Peace Science* 30 (2): 99–120.
- Downs, George W., David M. Roake, and Peter N. Barsoom. 1996. Is the Good News about Compliance Good News about Cooperation? *International Organization* 50 (3):379–406.
- Eck, Kristine, and Lisa Hultman. 2007. Violence against Civilians in War. *Journal of Peace Research* 44 (2):233–46.
- Fariss, Christopher J. 2013. Uncertain Events: A Dynamic Latent Variable Model of Human Rights Respect and Government Killing with Binary, Ordered, and Count Outcomes. Presented at the Annual Meeting of the Society for Political Methodology, University of Virginia, July.
- . 2014. Respect for Human Rights Has Improved over Time: Modeling the Changing Standard of Accountability in Human Rights Documents. *American Political Science Review* 108 (2):297–318.
- Fariss, Christopher J., and Keith Schnakenberg. 2014. Measuring Mutual Dependence between State Repressive Actions. *Journal of Conflict Resolution* 58 (6):1003–32.
- Fariss, Christopher. 2015. Replication Data for: The Changing Standard of Accountability and the Positive Relationship between Human Rights Treaty Ratification and Compliance. <http://dx.doi.org/10.7910/DVN/T177ZP>, Harvard Dataverse, V2.
- Gelman, Andrew, and Donald B. Rubin. 1992. Inference from Iterative Simulation Using Multiple Sequences. *Statistical Science* 7:457–511.
- Geweke, John. 1992. Evaluating the Accuracy of Sampling-Based Approaches to Calculating Posterior Moments. Pp. 169–93 in *Bayesian Statistics 4*, edited by J. M. Bernardo, J. Berger, A. P. Dawid and J. F. M. Smith. Oxford: Oxford University Press.
- Gibney, Mark, Linda Cornett, Reed M. Wood, and Peter Haschke. 2015. Political Terror Scale. Available from <http://www.politicalterroryscale.org> (last accessed: 21 October 2015).
- Gill, Jeff. 2008. *Bayesian Methods: A Social and Behavioral Sciences Approach*. New York: Chapman & Hall/CRC Statistics in the Social and Behavioral Sciences, 2nd edn. Chapman and Hall/CRC.
- Gleditsch, Kristian Skrede. 2002. Expanded Trade and GDP Data. *Journal of Conflict Resolution* 46 (5): 712–724.
- Goodman, Ryan, and Derek Jinks. 2003. Measuring the Effects of Human Rights Treaties. *European Journal of International Law* 14 (1):171–83.
- Hafner-Burton, Emilie M. 2013. *Making Human Rights a Reality*. Princeton, N.J.: Princeton University Press.
- . 2014. A Social Science of Human Rights. *Journal of Peace Research* 51 (2):273–86.
- Hafner-Burton, Emilie M., and James Ron. 2009. Seeing Double Human Rights Impact through Qualitative and Quantitative Eyes. *World Politics* 61 (2):360–401.
- Hafner-Burton, Emilie M., and Kiyoteru Tsutsui. 2005. Human Rights in a Globalizing World: The Paradox of Empty Promises. *American Journal of Sociology* 110 (5):1373–411.
- . 2007. Justice Lost! The Failure of International Human Rights Law to Matter Where Needed Most. *Journal of Peace Research* 44 (4):407–25.
- Hafner-Burton, Emilie M., Kiyoteru Tsutsui, and John W. Meyer. 2008. International Human Rights Law and the Politics of Legitimation – Repressive States and Human Rights Treaties. *International Sociology* 23 (1):115–41.
- Harff, Barbara. 2003. No Lessons Learned from the Holocaust? Assessing Risks of Genocide and Political Mass Murder since 1955. *American Political Science Review* 97 (1):57–73.

- Harff, Barbara, and Ted R. Gurr. 1988. Toward Empirical Theory of Genocides and Politicides: Identification and Measurement of Cases since 1945. *International Studies Quarterly* 32 (3):359–71.
- Hathaway, Oona A. 2002. Do Human Rights Treaties Make a Difference? *Yale Law Journal* 111 (8):1935–2042.
- Heidelberger, Philip, and Peter D. Welch. 1981. A Spectral Method for Confidence Interval Generation and Run Length Control in Simulations. *Communications of the ACM* 24:233–45.
- . 1983. Simulation Run Length Control in the Presence of an Initial Transient. *Operations Research* 31 (6):1109–44.
- Hill, Daniel W. Jr. 2010. Estimating the Effects of Human Rights Treaties on State Behavior. *Journal of Politics* 72 (4):1161–174.
- Hill, Daniel W. Jr., Will H. Moore, and Bumba Mukherjee. 2013. Information Politics v Organizational Incentives: When are Amnesty International's 'Naming and Shaming' Reports Biased? *International Studies Quarterly* 57 (2):219–32.
- Hill, Daniel W. Jr., and Zachary M. Jones. 2014. An Empirical Evaluation of Explanations for State Repression. *American Political Science Review* 108 (3):661–87.
- Hollyer, James R., and B. Peter Rosendorff. 2011. Why Do Authoritarian Regimes Sign the Convention Against Torture? Signaling, Domestic Politics and Non-Compliance. *Quarterly Journal of Political Science* 6:275–327.
- Hopgood, Stephen. 2006. *Keepers of the Flame: Understanding Amnesty International*. Ithaca, N.Y: Cornell University Press.
- . 2013. *The Endtimes of Human Rights*. Ithaca, N.Y: Cornell University Press.
- Jackman, Simon. 2008. Measurement. In *The Oxford Handbook of Political Methodology*, edited by Janet M. Box-Steffensmeier, Henry E. Brady and David Collier. Oxford: Oxford University Press, doi: <http://dx.doi.org/10.1093/oxfordhb/9780199286546.003.0006>
- . 2009. *Bayesian Analysis for the Social Sciences*. New York: JohnWiley and Sons.
- Jessee, Stephen A. 2015. 'Don't Know' Responses, Personality and the Measurement of Political Knowledge. *Political Science Research and Methods* (doi: <http://dx.doi.org/10.1017/psrm.2015.23>). Published online by Cambridge University Press: 19 June 2015.
- Keck, Margaret, and Kathryn Sikkink. 1998. *Activists beyond Borders: Advocacy Networks in International Politics*. Ithaca, N.Y: Cornell University Press.
- Keith, Linda Camp. 1999. The United Nations International Covenant on Civil and Political Rights: Does It Make a Difference in Human Rights Behavior? *Journal of Peace Research* 36 (1):95–118.
- . 2002. Constitutional Provisions for Individual Human Rights (1977–1996): Are They More than Mere 'Window Dressing'? *Political Research Quarterly* 55 (1):111–43.
- . 2012. *Political Repression Courts and the Law*. Philadelphia: University of Pennsylvania Press.
- Keith, Linda Camp, C. Neal Tate, and Steven C. Poe. 2009. Is the Law a Mere Parchment Barrier to Human Rights Abuse? *Journal of Politics* 71 (1):644–60.
- Keith, Linda Camp, and Steven C. Poe. 2004. Are Constitutional State of Emergency Clauses Effective? An Empirical Exploration. *Human Rights Quarterly* 26 (4):1071–97.
- Kim, Hunjoon, and Kathryn Sikkink. 2010. Explaining the Deterrence Effect of Human Rights Prosecutions for Transitional Countries. *International Studies Quarterly* 54 (4):939–63.
- Korey, William. 2001. *NGOs and the Universal Declaration of Human Rights: A Curious Grapevine*. Basingstoke: Palgrave Macmillan.
- Landman, Todd. 2005. The Political Science of Human Rights. *British Journal of Political Science* 35 (3): 549–72.
- Lord, Frederic M. 1980. *Applications of Item Response Theory to Practical Testing Problems*. Mahwah, N.J: Erlbaum Associates.
- Lord, Frederic M., and Melvin R. Novick. 1968. *Statistical Theories of Mental Test Scores*. Boston. Mass: Addison-Wesley.
- Lupu, Yonatan. 2013a. Best Evidence: The Role of Information in Domestic Judicial Enforcement of International Human Rights Agreements. *International Organization* 67 (3):469–503.
- . 2013b. The Informative Power of Treaty Commitment: Using the Spatial Model to Address Selection Effects. *American Journal of Political Science* 57 (4):912–25.

- . 2015. Legislative Veto Players and the Effects of International Human Rights Agreements. *American Journal of Political Science* 59 (3):578–94.
- Marchesi, Bridget, and Kathryn Sikkink. 2015. The Effectiveness of the International Human Rights Legal Regime: What Do We Know and How Do We Know It? (Harvard University working paper).
- Marshall, Monty, Keith Jagers, and Ted R. Gurr. 2013. Polity IV Project: Political Regime Characteristics and Transitions 1800–2013 Dataset Users' Manual. Available from [www.systemicpeace.org/polity/polity4.htm](http://www.systemicpeace.org/polity/polity4.htm) (last accessed: 21 October 2015).
- Martin, Andrew D., and Keven M. Quinn. 2002. Dynamic Ideal Point Estimation via Markov Chain Monte Carlo for the U.S. Supreme Court, 1953–1999. *Political Analysis* 10 (2):134–53.
- Mayerfeld, Jamie. 2016. *The Architecture of Human Rights: Why Constitutional Government Requires International Human Rights Law*. Philadelphia: University of Pennsylvania Press.
- Morrow, James D. 2007. When Do States Follow the Laws of War? *American Political Science Review* 101 (3):559–72.
- Moyn, Samuel. 2010. *The Last Utopia: Human Rights in History*. Cambridge, Mass: The Belknap Press of Harvard University Press.
- Murdie, Amanda, and David R. Davis. 2012. Looking in the Mirror: Comparing INGO Networks across Issue Areas. *Review of International Organizations* 7 (2):177–202.
- Murdie, Amanda, and Tavishi Bhasin. 2011. Aiding and Abetting? Human Rights INGOs and Domestic Anti-Government Protest. *Journal of Conflict Resolution* 55 (2):163–91.
- Neumayer, Eric. 2005. Do International Human Rights Treaties Improve Respect for Human Rights? *Journal of Conflict Resolution* 49 (6):925–53.
- Nordås, Ragnhild, and Christian Davenport. 2013. Fight the Youth: Youth Bulges and State Repression. *American Journal of Political Science* 57 (4):926–40.
- Pemstein, Daniel, Stephen A. Meserve, and James Melton. 2010. Democratic Compromise: A Latent Variable Analysis of Ten Measures of Regime Type. *Political Analysis* 18 (4):426–49.
- Plummer, Martyn. 2010. JAGS (Just Another Gibbs Sampler) 1.0.3 Universal. Available from <http://mcmc-jags.sourceforge.net/> (last accessed: 21 October 2015).
- Poe, Steven C., and C. Neal Tate. 1994. Repression of Human Rights to Personal Integrity in the 1980s: A Global Analysis. *American Political Science Review* 88 (4):853–72.
- Poe, Steven C., C. Neal Tate, and Linda Camp Keith. 1999. Repression of the Human Right to Personal Integrity Revisited: A Global Cross-National Study Covering the Years 1976–1993. *International Studies Quarterly* 43 (2):291–313.
- Poole, Keith T., and Howard Rosenthal. 1997. *A Political-Economic History of Roll Call Voting*. New York: Oxford University Press.
- Posner, Eric A. 2014. *The Twilight of Human Rights Law*. Oxford: Oxford University Press.
- Powell, Emilia Justyna, and Jefferey K. Staton. 2009. Domestic Judicial Institutions and Human Rights Treaty Violation. *International Studies Quarterly* 53 (1):149–74.
- Quinn, Keven M. 2004. Bayesian Factor Analysis for Mixed Ordinal and Continuous Responses. *Political Analysis* 12 (4):338–53.
- Rasch, Georg. 1980. *Probabilistic Models for Some Intelligence and Attainment Tests*. Chicago: The University of Chicago Press.
- Roth, Brad. 2001. Understanding the Understanding: Federalism Constraints on Human Rights Implementation. *Wayne Law Review* 47:891–907.
- Rummel, Rudolph J. 1994. Power, Genocide and Mass Murder. *Journal of Peace Research* 31 (1):1–10.
- . 1995. Democracy, Power, Genocide, and Mass Murder. *Journal of Conflict Resolution* 39 (1):3–26.
- Sandholtz, Wayne. 2012. Treaties, Constitutions, Courts, and Human Rights. *Journal of Human Rights* 11 (1): 17–32.
- Schnakenberg, Keith E., and Christopher J. Fariss. 2014. Dynamic Patterns of Human Rights Practices. *Political Science Research and Methods* 2 (1):1–31.
- Shadish, William R. 2010. Campbell and Rubin: A Primer and Comparison of Their Approaches to Causal Inference in Field Settings. *Psychological Methods* 12 (1):3–17.

- Shadish, William R., Thomas D. Cook, and Donald T. Campbell. 2001. *Experimental and Quasi-Experimental Designs for Generalized Causal Inference*. Belmont, Calif: Wadsworth Publishing.
- Sikkink, Kathryn. 2011. *The Justice Cascade: How Human Rights Prosecutions Are Changing World Politics*. New York: Norton Series in World Politics.
- Simmons, Beth A. 2000. International Law and State Behavior: Commitment and Compliance in International Monetary Affairs. *American Political Science Review* 94 (4):819–35.
- . 2009. *Mobilizing for Human Rights: International Law in Domestic Politics*. Cambridge: Cambridge University Press.
- Simmons, Beth A., and Daniel J. Hopkins. 2005. The Constraining Power of International Treaties: Theory and Methods. *American Political Science Review* 99 (4):623–31.
- Sinkovits, Robert S., Pietro Cicotti, Shawn Strande, Mahidhar Tatineni, Paul Rodriguez, Nicole Wolter, and Natasha Bala. 2011. Data Intensive Analysis on the Gordon High Performance Data and Compute System. KDD’11 Proceedings of the 17th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pp. 747–48. doi:10.1145/2020408.2020526
- Smith-Cannoy, Heather. 2012. *Insincere Commitments: Human Rights Treaties, Abusive States, and Citizen Activism*. Washington, D.C: Georgetown University Press.
- Taylor, Charles Lewis, and David A. Jodice, eds. 1983. *World Handbook of Political and Social Indicators*, 3rd edn. Vol. 2, Political Protest and Government Change. New Haven, Conn: Yale University Press.
- Treier, Shawn, and Simon Jackman. 2008. Democracy as a Latent Variable. *American Journal of Political Science* 52 (1):201–17.
- Trochim, William M.K., and James P. Donnelly, eds. 2008. *Research Methods Knowledge Base*, 3rd edn, Mason, Ohio: Atomic Dog.
- Voeten, Erik. 2000. Clashes in the Assembly. *International Organization* 54 (2):185–215.
- Von Stein, Jana. 2005. Do Treaties Constrain or Screen? Selection Bias and Treaty Compliance. *American Political Science Review* 99 (4):611–22.
- Vreeland, James Raymond. 2008. Political Institutions and Human Rights: Why Dictatorships Enter into the United Nations Convention Against Torture. *International Organization* 62 (1):65–101.
- Western, Bruce. 1999. Bayesian Methods for Sociologists: An Introduction. *Sociological Methods & Research* 28 (1):7–34.
- Wong, Wendy H. 2012. *Internal Affairs: How the Structure of NGOs Transforms Human Rights*. Ithaca, N.Y: Cornell University Press.