Hauser Globalization Colloquium Fall 2010

Professors Ryan Goodman & Robert Keohane

Furman Hall 900, Pollack Colloquium Room
Wednesdays 2:00 pm-3:50pm
(unless otherwise noted)

Schedule of Sessions (subject to modification)

September 15   **Professor Eric Posner**
Human Rights, the Laws of War, and Reciprocity

September 22   **Professor Michael Doyle**
A Global Constitution? The Struggle over the UN Charter

October 6   **Professor Mary Dudziak**
Law, War, and the History of Time

October 13   **Professor Tim Buthe**
Standards for global markets: domestic and international institutions for setting international product standards

October 20   **Professor Kal Raustiala**
Information and International Agreements
Background Readings:
Police Patrols and Fire Alarms in the NAAEC
The Rational Design of International Institutions

October 22
(Friday)   **Professor Peter Katzenstein**
The Transnational Spread of American Law: Legalization as Soft Power

November 10   **Professors Oona Hathaway & Scott Shapiro**
Outcasting: Enforcement in Domestic and International Law

November 17   **Professors Ann Marie Clark & Kathryn Sikkink**
"Information Effects and Human Rights Data: Is the Good News about Increased Human Rights Information Bad News for Human Rights Measures?"
Background Reading: Emilie M. Hafner-Burton, & James Ron, Seeing Double: Human Rights Impact Through Qualitative and Quantitative Eyes, World Politics, 2009.

December 1   **Professors Kevin Davis and Benedict Kingsbury**
Obligation Overload: Restructuring the Obligations of Fragile or Failed States

December 3
(Friday)   **Professor Beth Simmons**
Subjective Frames and Rational Choice: Transnational Crime and the Case of Human Trafficking

Information Effects and Human Rights Data:
Is the Good News about Increased Human Rights Information
Bad News for Human Rights Measures?[1]

Ann Marie Clark   Kathryn Sikkink
Purdue University   University of Minnesota
clarkam@purdue.edu   sikkink@umn.edu

November 2010 draft

Comments are welcome.
Please check with the authors before citing.

Introduction

In a recent book review essay Hafner-Burton and Ron point to what they call the "persistent gap" between the rhetorical success of human rights, and the "empirical reality" of lagging actual improvement of human rights practices on the ground (Hafner-Burton and Ron 2009:13). Hafner-Burton and Ron imply that qualitative researchers have been too optimistic about the results of human rights policies, while quantitative researchers have been more pessimistic, and that such differences of interpretation are shaped "by the choice of research method." When it comes to isolating causal indicators of human rights improvement, the essay attributes the sometimes discouraging results of selected large-*n* cross-national studies to a better purchase on reality. We find Hafner-Burton and Ron's characterization of the debate between overly optimistic qualitative researchers, and the implicitly more accurate quantitative researchers a potentially misleading way to approach these issues. We seek to focus the discussion more constructively, by demonstrating that the choice of data, the artifacts of data gathering, and the information environment within which these take place, *which are very different from the choice of method*, may instead be coloring conclusions. Both qualitative and quantitative research findings are influenced by such choices, but the problem for quantitative human rights research is particularly acute because it relies heavily on two major data bases coded from the same two sources to summarize human rights change in the international system. Specifically we suggest that systematic "information effects" in these data sets may contribute to the pessimistic findings of the quantitative literature. We consider these quantitative measures valuable resources, we have used them in our own work, and we do not intend to launch a blanket criticism of the human rights measures. Rather, we raise these issues in the hope of contributing constructively to a better causal understanding of human rights change.

For example, there has been large debate prompted by quantitative literature on the impact of human rights treaty ratification on human rights practices in countries around the world. (See Keith 1999; Hathaway 2002; Weissbrodt 2003; Hafner-Burton and Tsutsui 2005; Landman 2005; Neumayer 2005; Simmons 2009.) Goodman and Jinks (2003) have recognized that "this critical question is empirical in character," and fundamentally linked to the quality of the data used. In particular, they criticized work by

Hathaway (2002) for failing to recognize the ways in which treaty ratification might heighten scrutiny or the provision of information in ways that might make it appear as if the human rights situation has worsened as a result of ratification. The argument we make here is that such information effects are not limited to research involving treaty ratification, but might be a more systematic feature of human rights measures as a result of increasing amount of human rights information available over time.

We believe that the question of which methodological approach is inherently more accurate is unproductive and probably unanswerable. Aside from the fact that we could name many pessimistic case study researchers and many optimistic users of large-n data, the 2009 review essay left out or marginalized the importance of several large-*n* studies that do find evidence of improvement as a result of international human rights treaty mechanisms. (See, among others, Landman (2005); Neumayer (2005); and the comprehensive study of the impact of human rights law by Simmons (2009), many of whose findings were in circulation prior to publication.) Nevertheless, the important debate over how to interpret findings should not be replaced by or diminished as "difference of opinion." There need not be a dichotomy between qualitative and quantitative research on human rights. Some recent scholarship has engaged the measurement issues in a more sustained way that does not provoke a divide (Landman and Carvalho 2010). Our main point is that, far from pitting qualitative and quantitative analysts against one another or characterizing either as optimistic or pessimistic about human rights improvement, the force of argument rests on good evidence. Human rights data is one kind of evidence, but it is only as good as its sources, and human rights information has increased over time, so that all analysts should maintain a strong awareness of the limits of inference based on cross-national data.

In a separate exchange, the custodians of the two data sets, the Political Terror Scale (PTS) and the Cingranelli-Richards Human Rights Dataset (CIRI), have clashed over coding approaches to the annual human rights reports of Amnesty International and U.S. Department of State Department, on which the two indices rely (Cingranelli and Richards 2010; Wood and Gibney 2010). These recent pieces call for further specific attention to issues of human rights data and measurement and awareness about the sources giving rise to our data.

Recent human rights source reports from AI and the U.S. State Department typically contain much more and better information than earlier ones. More recent reports also tend to document a wider range of human rights violations. For example, over time, the State Department reports have expanded to include acts of terrorism, workers rights, and women's rights. An increase in the quality and quantity of information about human rights violations in the world and greater attention to the full range of human rights is good news for scholars and practitioners in this area, but it could have some unanticipated implications for the general validity of human rights measures like CIRI and PTS and the explanatory power of research using solely those measures – unless ways can be found to correct for deficiencies in the data. In this paper, we present and begin to test some hypotheses about information effects, defined as the effects of the varying availability and precision of country-specific information on human rights that has accompanied the rise in prominence of human rights concerns in international politics in recent decades. We concentrate on the PTS and CIRI data sets and on the two annual reports from which they are coded, although some information effects are likely to apply broadly to all uses of human rights data. We find that the PTS and CIRI indices appear to contain some embedded information effects related to the increasing availability of human rights information over time. One conclusion that follows from our analysis is that it may be more difficult to mark real human rights improvements than human rights deterioration using these datasets.

To permit closer investigation of variation in reports over time and comparison with other sources of data, we have chosen to focus on the Latin American region, a region with whose human rights profiles we are most familiar. The narrowed focus enables us to look further back in time at the ways in which reports may have changed in length and scope for countries in this particular region. Arguably, given the uncertain and often conflicting nature of contemporaneous reports of human rights abuse, it would be useful to reconstruct the scales for a sample of countries, given what is now known, and compare them.[2] This is not possible in this paper, but we do employ comparative data from five different truth commission reports and one supplemental report for the countries of Argentina, Brazil, Chile, Guatemala,

---

[2] We thank Robert Keohane for a suggestion that prompted our thinking on this.

and Peru. These reports provide annual data on deaths and disappearances and allow an initial exploration of hypotheses that might eventually be tested more fully.

We find evidence of a mild-to-moderate information effects. We suggest potential vulnerabilities of the two indices and test how troublesome such vulnerabilities might be. We find that, first, the PTS and CIRI indices both are correlated to some degree with report length, when we control for variation related to systematic levels of abuse. Second, we illustrate how the expansion of the kinds of abuses covered by the reports themselves has an effect that users of the data should be aware of: in the case of Brazil, which we examine in detail, coding levels of later cases of widespread police abuses, for example, rival the coding of early large-scale disappearances and deaths. Users of the data should not assume that repression in two similarly coded periods was similar. To get a more general feel for how well coding corresponds with the worst physical integrity abuses, which they are intended to measure, we then compare PTS and CIRI coding with the numbers of deaths and disappearances documented by truth commission reports. We find that codes remain "sticky" at high levels of human rights abuses, meaning that the worst coding levels may be likely to persist over time for severe cases of abuse. If so, then the scales will be less likely to register improvements than they are to show deterioration in the human rights environment of a country. This "stickiness" and "substitutability" are two features that Hafner-Burton and Ron (2009) mention in passing, but we come to different conclusions about the overall implications for understanding the dynamics of human rights change. Ironically, some of the key cases that provoked the development of worldwide human rights advocacy and active domestic human rights NGOs – cases such as Argentina and Chile, with early, severe, and large scale repression in the form of deaths and disappearances – might be the cases for which bias caused by information effects most drown out any improvements.

Information effects

Since the mid-1970s, the human rights realm has passed from being relatively information-poor to being much more information-rich. It is always difficult to collect information on human rights violations. In 1975, there were very few sources of reliable human rights information in the world. Even

now, governments often try to hide such information, downplay it, or dismiss it, so we can rarely rely on states alone to provide reliable data about their own human rights practices (Innes de Neufville 1986:698). The main contributors to the information-richer environment have been NGOs, IOs, and a few states that systematically gather and publish cross-national human rights information. Since 1975, the number of NGOs and IOs reporting on human rights has proliferated. Violations are less likely to be hidden and unknown than they were before, and researchers have more sources to draw on. This is a positive development for knowledge about human rights, but it can provoke information effects in the data, a phenomenon Keck and Sikkink referred to more generally as the "information paradox": as NGOs draw attention to new forms of human rights violations and make more information available about human rights, in some cases it may appear that the human rights situation is getting worse when in fact we simply have more information about human rights practices (Keck and Sikkink 1998:194).

We reserve the term "human rights information paradox" to describe a situation where international organizations, governments and human rights NGOs, by producing increasingly more information about a wider range of human rights violations, may give the impression that a human rights situation is worsening. Such a paradox influences the information environment in which both the qualitative and quantitative researcher work, and thus may have an influence on their evaluation of human rights change. We will use the term "information effects" to refer to the effects of this varying availability and precision of country-specific information on human rights, especially with regard to the two major data sources and two major data bases used in much quantitative human rights research.

Sources and Coding: How PTS and CIRI Use Human Rights Reports

As mentioned above, many statistical studies of international human rights rely on one of two prominent measures of states' human rights performance, the Political Terror Scale (PTS) and the Cingranelli-Richards Human Rights Data Set (CIRI). These two measures in turn rely on two annual reports, the *Amnesty International Annual Report*, and the U.S. State Department annual *Country Reports on Human Rights Practices*. PTS and CIRI cover most countries of the world and offer scholars data that can be used to test causal hypotheses about human rights in a cross-national setting and over time

(Cingranelli and Richards 2008; Gibney, Cornett et al. 2008). Both PTS and CIRI provide numeric indicators based on yearly coding of the texts of Amnesty International's annual reports and the U.S. State Department's annual reports on human rights practices, which have been issued regularly for several decades.

While CIRI coding begins in 1981, PTS data have recently been posted to incorporate the coding that Poe et al. (1999) extended back through 1976. The CIRI and PTS coding schemes are described in detail in many places, including the scales' own web sites, so we will only briefly describe them here.

The Political Terror Scale (PTS). The PTS data set is a standards-based scale, in the sense that its coding categories relate how widespread political violence, defined as threats to physical integrity rights, is in a society for a given year. PTS characterizes the content of Amnesty International and U.S. State Department annual entries in their annual country reports on a scale of 1 to 5, pertaining to how broadly state-based political violence, defined as threats to physical integrity rights, extends to the country's population. A coding of 1 represents the prevailing rule of law, with no or very rare incidence of political violence, while 5 represents extended political violence in which "terror has extended to the whole population" (Gibney, Cornett et al. 2008). The PTS does not attempt to count incidents of repression.

Cingranelli Richards Human Rights Data Set (CIRI). The CIRI data set creates a score for each of four types of physical integrity abuses: political imprisonment, torture, political killing, and disappearances. CIRI's Physical Integrity Index adds the four scores together. It is not a count of violations, but the early coder guidelines suggested that coders assign values based on whether violations are "practiced frequently," "practiced occasionally," or "have not occurred," and, where possible, the coding should be based on numbers if reported (Cingranelli and Richards 2008:7). The coding ranges from 0 to 2 for each type of violation, assigning 2 for no violations of a particular type, 1 for 0-49 violations of that type, and 0 for 50 or more violations of a particular type. Thus, the Physical Integrity Index ranges from 0 to 8. In practice, because there often are no specific estimates of violations in the reports, the codes assigned may reflect coders' qualitative estimates of the level of violations based on the text of the reports (Cingranelli and Richards 2008:8), and Cingranelli and Richards have recently

emphasized that this is the typical way that the reports are coded (Cingranelli and Richards 2010). For example, the coder guide suggests that language indicating a pattern of abuses, as opposed to isolated incidents, often "merit[s] a 0" (Cingranelli and Richards 2008:8). The CIRI physical integrity scale represents the added values of each type of violation for a particular year. Unlike the PTS, CIRI does not have a separate AI and State Department coding for each entry; instead, it looks to the State Department coding as its primary source and uses the AI report as a second source for physical integrity rights (Cingranelli and Richards 2008:3).

Wood and Gibney (Wood and Gibney 2010) conducted a detailed comparison of the two scales that has been published with a rejoinder from Cingranelli and Richards (2010). In brief, Wood and Gibney point out key differences that result from comparing the standards-based of PTS and the composite, event count-style coding of CIRI's physical integrity scale. Cingranelli and Richards emphasized that CIRI of necessity does rely on coder judgments based on the report texts rather than numbers, and that it, too, seeks to compare reported human rights conditions with particular standards. However, the coding approaches are somewhat divergent although they rely on similar sources. Instead of reviewing the potentially significant debate about the two scales in detail, we want to look more broadly at information effects. Therefore, below we focus on the characteristics of the data that are most relevant for the analysis of information effects as they might be reflected in the data sets. As outlined in Table 1 and below, we hypothesize that information effects, if they exist, will manifest themselves differently in the different data sets.

The textual sources for PTS and CIRI

One issue not highlighted in the recent exchange about PTS and CIRI is that coding is based on primary source texts. Remembering that fact helps to put the data in context. Landman and colleagues have noted in both academic research and research directed toward the policy community that there often is a "lack of awareness" about the sources of data (Landman and Häusermann 2003:17; Landman and Carvalho 2010). This does not solely affect human rights studies, but as Landman and Häusermann have pointed out, lack of awareness also has tended to result in "over-reliance on a few source materials,"

which include the AI and U.S. State Department sources (Landman and Häusermann 2003:17). A bit of attention to how the reports are produced is warranted here.

The reports generated by AI and the U.S. Department of State are produced contemporaneously, with all of the associated advantages and disadvantages. First, the advantages: over and over again, reports of outside monitors, however imperfect, have come much closer to revealing the nature of repression than the self-reporting of states. For researchers, the value of single source texts as a data resource, issued annually in similar format, cannot be overstated.   We are even more fortunate that two such series exist. Still, we have to live with, and not ignore, certain disadvantages.

First, the reports are compiled by human actors who are often distant from the places where the violations are occurring. AI reports are prepared in London. They draw on material gathered from domestic human rights organizations and sometimes from on-site missions. But in situations of serious human rights violations, governments often prohibit on-site missions and repress domestic human rights organizations. Secondly, investigators of contemporaneous violations are hindered by repressive apparatuses that do everything possible to hide their actions, and even well meaning actors doing the best job they can under the circumstances will produce imperfect reports. Finally, in the midst of major episodes of human rights violations, it is often very difficult to pinpoint and identify exactly what is going on and how many victims there are of different types of human rights violations. Widespread repression also blocks domestic NGOs' capacity to link with outside organizations and, sometimes, their ability to exist and even to function. As such, we do not believe that the actors who produce the reports can do much more to improve such contemporaneous reporting. It is the result of the nature of the enterprise.

It is still worth noting that State Department and Amnesty International annual human rights reports cannot be perfectly objective measures of human rights practices, nor do the organizations claim that the texts are comprehensive. The State Department has political goals that may affect human rights reporting, as we discuss in the Guatemala case. Amnesty International is committed to a human rights ethos that may make it difficult to characterize some reduction in numbers as "improvement" in the context of serious on-going violations of human rights. These kinds of biases, where they exist, are likely

to enter the coded data, since the databases emphasize that coding should be consistent with the textual content of the country reports and not attempt to filter the information.[3] Even with these caveats, these are extremely important reports and data sets. But researchers who use the data sets should be constantly aware that they are not perfect measures of human rights, and that the organizations themselves do not construct their reports as a basis for quantitative indices.

Turning to the practical production of the texts themselves, and their relevance for the possible detection of information effects in the data, it should also be recognized that the reporting organizations have not necessarily followed the same formula over the years, nor has the level of information remained constant. This is especially apparent when more recent annual reports are compared with early ones. First, the early reports covered fewer countries. When Poe, Tate, and Keith expanded the PTS back from 1980 to 1976, they explicitly recognized that AI tended to leave out reports for countries with less problematic human rights records and, when reports began, the U.S. State Department only reported on countries receiving U.S. aid (Poe, Tate et al. 1999:300). Second, as mentioned above, later reports discuss more forms of rights violations. Third, and perhaps least apparent to potential data set users, the later reports very likely contain comparatively more and better information, other things being equal.

Let us expand on this last point. One reason the information is more plentiful and better is that both Amnesty International and the U.S. State Department have greatly increased their organizational research and reporting capabilities over the last thirty years. In the early 1970s, before the State Department was mandated to produce the human rights country reports, there was one person in the State Department with a responsibility for human rights. By the end of the Carter administration, there were 20 professional human rights staff positions in the State Department (Griesgraber 1983:103,106). By the end of the Clinton Administration, the Human Rights Bureau in the State Department had more than 100 staff members (Sikkink 2004:206). Amnesty International staff size grew from 40 in 1976 to 150 in

---

[3] Ron et al. found that Amnesty International and Human Rights Watch, two of the largest international human rights organizations, also do some strategic targeting in their issuance of press releases and special reports (Ron, J., H. Ramos, et al. (2005). "Transnational Information Politics: NGO Human Rights Reporting, 1986-2000." International Studies Quarterly **49**(3): 557.) However, we would expect any effects of such targeting to be diminished in the context of the similarly formatted annual reports.

1981, after its growing influence was recognized by the Nobel Peace Prize in 1977 (Dezalay and Garth 2006:239, citing Korey [1998]:179). In other words, the bureaucratic and research capacity of both AI and the State department to report on human rights increased dramatically in the period over which they have issued their annual reports.

Another reason for more plentiful information is that the State Department and Amnesty International do not rely solely on their own staff to produce the human rights reports. They also rely on other organizations, especially domestic NGOs, with specialized human rights knowledge. After host governments, human rights groups were the most important source for human rights information used by embassies to prepare the State Department's human rights reports, according to a 1990 US GAO account.[4] The sheer number of such groups, and consequently the amount of information available, has increased broadly since the beginning of the time series covered by PTS and CIRI. To illustrate, the total number of international human rights NGOs doubled between 1983 and 1993, and their budgets and staffs grew dramatically. The network grew in Latin America as well. In the 1970s and 1980s, domestic human rights organizations appeared throughout Latin America, increasing from 220 to 550 between 1981 and 1990.[5] Their growing capacity allowed human rights NGOs to do more on-site investigations and missions, also enhancing the quality of their reporting with primary research. In addition, as human rights organizations grew, their members and constituencies demanded that they address a wider range of human rights issues. Amnesty International, for example, had started as an organization with a relatively narrow mandate of working on prisoners of conscience, but over the years expanded its mandate to take on a much wider range of human rights (Baehr 1994; Clark 2001; Hopgood 2006).

Hypotheses

---

[4] U.S. General Accounting Office (GAO) (1990). Human Rights: State Department's Commitment to Accurate Reporting Has Increased. Washington, DC, Report to the Chairman, Legislation and National Security Subcommittee, Committee on Government Operations, House of Representatives, September 1990. p. 22.

[5] Data from Wiseberg, L. S. and H. M. Scoble, Eds. (1981). Human Rights Directory: Latin America, Africa, and Asia. Washington, DC, Human Rights Internet. and Human Rights Internet Reporter (1990). Human Rights Directory: Latin America and the Caribbean, HRI. **13**.

As a result of the general reporting changes described above, we suggest that some set of "information effects" resulting from changes in reporting is likely. Our expectations are summarized in Table 1, and we elaborate on them here.

First, if sheer amount of information drives harsher coding, we ought to observe some correlation of coding with report length, controlling for actual change in repression (see row 1 of the Table 1). Here, the source of the error would lie with the source text. The PTS does not count numbers of violations. While this may limit the scale's ability to be specific about amount and types of violations, if the reports' descriptive characterizations of human rights abuses are similar over time, then on the face of it we might expect standards-based coding to be less subject to an information effect caused simply by more detailed reports. On the other hand, because the CIRI index has some (at least implicit) count criteria, we might expect it to be somewhat more susceptible to information effects based on amount of information or level of detail.

Second, if the quality of reporting is improving over time so that reports reflect more detail on type of political violence (regardless of report length), errors in comparability could be introduced. This sort of error could lie with the source text, if all countries do not have equal levels of transparency and so reports cannot all offer the same level of detail either cross-nationally or over time. (See row 2 of the table.) Or, it could lie with measurement: We might expect the scales to differ in their ability to differentiate changes in type, scope, and intensity of repression either over time or cross-nationally. (See row 3.) Without extended in-depth textual coding, we are unable to perform a comprehensive analysis of whether and how the types of violations reported have varied, but we do examine two cases in depth to probe the plausibility of this dynamic: Brazil and Guatemala. While we expect the PTS to be stronger in its resistance to "noise" in reporting length and detail, the PTS's "weakness" might lie in lack of ability to give coders a way to differentiate among in types of violations that would be coded in the aggregate at similar levels. On this score, CIRI's attempt to disaggregate types of violations might be a strength. Further, the PTS criteria for the worst violations include how widespread violations are in the extent to which they affect the population, and this differs from CIRI's individual component scores, which in fact,

do not attempt to address scope. CIRI will assign a poor rating to a higher levels of any given particular

kind of violation, independent of whether or not it affects a large portion of the population.[6] This might

mean that CIRI would flag some violations that most would agree were very serious, even if they did not

extend to broad sectors of the population (and, it might be difficult to judge how broad some kinds of

violations were). However, the most intense uses of just a single type of type of abusive practice (or

fewer than all four types) might not be reflected as the poorest possible score in CIRI's physical integrity

index, which gives equal weight to all four types of abuses.

Third, error can be introduced if coding is not as sensitive to changes at higher levels of

repression as it is to low level changes. (See row 4). We hypothesize that this problem, a problem with

measurement, might show up more with CIRI, since 50 or more violations of the same kind receive the

same coding.   However, a measurement problem would also arise if coders (or texts' authors) read a

country's known previous repression forward. A reputational effect could occur, in other words.

A fourth kind of error may result if organizations themselves, for whatever reason, devote

differing levels of critical attention to particular countries in their annual reports (See row 5.) This could

happen as a result of political bias or simple variation in the level of resources or access the reporting

organizations were able to devote to country reporting. We are convinced by the analysis of Poe et al.

(Poe, Carey et al. 2001) and the account of Innes de Neufville mentioned above (Innes de Neufville 1986)

that this sort of bias is likely to have diminished over time, but it is stark for the Guatemala case, which

we examine in depth.

To explore these questions, we collected two kinds of additional data. First, to enable analysis of

report length, we performed word counts of AI reports and U.S. State Department reports by country

entry and year. Word counts for all countries included in the reports were relatively easy to come by for

those already in electronic format on the web, so we have comprehensive word counts for Amnesty

International reports from 1994 to 2006, and for the U.S. State Department from 1999 to 2006. For the

Latin American countries, we also electronically scanned the earlier reports back to 1975 (for Amnesty

---

[6] David Cingranelli, conversation with Ann Marie Clark, 15 February 2009.

International) and 1981 (State Department) to acquire a longer word-count series for the Latin American part of the sample.

Second, as an independent source of human rights data for comparison, we also employ the information compiled by the national truth commissions on deaths and disappearances for the five Latin American truth commissions, organizing it by year for each country to match the organization of the CIRI and PTS data.[7] The truth commissions categorized information on states' lethal repression of individuals in different ways: either as deaths and disappearances combined (Chile, Guatemala and Peru), just disappearances (Argentina), or just deaths (El Salvador). In Argentina, few people were killed directly before disappearance, so virtually all of the deaths documented during the period covered by the truth commission happened through disappearance. While these differences matter for the politics of remembering, for legal accountability, and for more detailed studies of the countries' repressive apparatus, most people who study disappearances agree in their understanding that virtually all of the disappeared people in each of these countries were killed. For these reasons and for this limited purpose we feel it is appropriate to consider a disappearance as equivalent to a death, in order to use all of the available truth commission data here.

In order to compare the scales in the regression analyses whose results are reported below, we have transformed the values for each to a zero-to-one scale following Wood and Gibney (2010), with zero representing the worst possible human rights score in either scale and 1 representing the best.[8]

Analysis

Our question could be phrased as asking whether more detailed reporting over the years (more information) may have introduced a negative bias into the databases. Since the PTS and CIRI coding is based on the texts of AI and U.S. State Department reports, it is conceivable that longer reports might produce more severe coding independent of the actual severity of abuses. The sheer availability of more

---

[7] The information from truth commission reports is an extension of the data in Sikkink, K. (2004). Mixed Signals: U.S. Human Rights Policy and Latin America. Ithaca, NY, Cornell.

[8] Thus, for the transformed PTS, a 1 corresponds to the original PTS value of 1 and 0 corresponds to an original PTS value of 5. For the transformed CIRI scale, a 1 corresponds to a CIRI physical integrity score of 8 and a 0 corresponds to a CIRI physical integrity score of 0.

information on human rights abuses might produce a worse human rights score, all other things being equal. However, the information effect has to be separated from a simple association of poor scores with longer reports because, presumably, states with truly worse human rights situations and numerous violations would require longer reports. To find evidence of an information effect, we would have to observe an association of longer reports with poorer human rights scores that is independent of actual human rights abuses.

It is also possible that less information from a country that is simply very closed, whatever the time period, could result in a misreading of the human rights situation there. At very high levels of repression, governments may simply eliminate all human rights monitors. More human rights information tends to be available for countries in an intermediate situation with relatively high levels of repression, but that are open enough to have larger numbers of human rights NGOs that document abuse.

To check whether a statistically significant growth trend in length of reports is apparent on average by country, we used pooled cross-national time series regression using fixed effects and robust standard errors to estimate:

(1) $\text{Wordcount}_{it} = \text{year}_{it} + \text{constant}_{it} + u_{it} + e_{it}$

This estimate would not show an information effect; it only tells us whether there is a pattern over time of longer reports on the whole. Results are shown in Table 2. The estimate shows no evidence of a general time trend in the length of AI reports on Latin America from 1976-2006 – report entries were about the same length on average over time. As we might guess from the lines on the graph in Figure 1, however, the length of State Department reports did increase substantially over time: the Latin America reports grew on average by almost 344 words per year in their coverage of the years 1981-2006. Globally from 1999-2006 (the years for which we have global word counts for both the State Department and AI), the State Department report entries for each country lengthened by an average of about 297 words per year,

while AI report entries shortened by a small but statistically significant average of 16 words per year.[9]

The mean length of State Department reports both globally and in Latin America peaked in 2001, as

depicted in Figure 1.

Even where a time trend is not evident in the yearly averages of report length, it is important to

ask whether coding is at all correlated with individual reports' variation in length separate from

worsening violations. As we mentioned, though, we also have to control for actual human rights levels.

As an imperfect way to control for "true" variation in human rights levels, we use the State Department

(SD) PTS score as a control variable when we test for investigate information effects based on AI reports,

and vice versa.   Although it would be ideal to regress the measure on independent information, we

believe this offers a rough control given data limitations. We included a lagged dependent variable to

control for the previous year's score. Thus, we regressed the human rights coding score for a given

country on the score of the previous year, that logged word count of the associated report for the current

year, and the coding score from the other source, using pooled cross-national time series regression with

fixed effects and robust standard errors.[10] There are three different scores at issue: the PTS scores based

on AI, the PTS scores based on the State department (SD), and CIRI, which uses the SD reports as the

primary referent. Thus, our estimates are as follows.

For the AI-*pts* score:

(2) $AI\_pts_{it} = AI\_pts_{i(t-1)} + \ln(AI\_wordcount_{it}) + SD\_pts_{it} + constant_{it} + u_i + e_{it}$

For the SD-*pts* score:

(3) $SD\_pts_{it} = SD\_pts_{i(t-1)} + \ln(SD\_wordcount_{it}) + AI\_pts_{it} + constant_{it} + u_i + e_{it}$

---

[9] This 1999-2006 selection is somewhat arbitrary; these are the years for which on-line reports were available from the State Department.

[10] Some researchers advise using ordered probit when analyzing the indices; similar models that were estimated using ordered probit and produced the same pattern of results and similar levels of statistical significance.

For the CIRI-*physical integrity* score, we regressed CIRI-*physint* on State Department word count and the AI-*pts* score, since CIRI's primary text basis is the State Department reports:

(4) $CIRI\_physint_{it} = CIRI\_physint\_pts_{i(t-1)} + \ln(SD\_wordcount_{it}) + AI\_pts_{it} + constant_{it} + u_i + e_{it}$

Since CIRI also uses AI reports to supplement its coding judgments, we also regressed the CIRI physical integrity score on AI word count, controlling for SD score:

(5) $CIRI\_physint_t = CIRI\_physint\_pts_{i(t-1)} + \ln(AI\_wordcount_{it}) + SD\_pts_t + constant_{it} + u_i + e_{it}$

where u_i refers to country-specific fixed effects.

The results are reported in Table 3, columns 1, 3, 5, and 7. We do find evidence of an information effect relative to report length. According to our results, the coefficient for length (the *wordcount* variable) is positive, meaning that on average, as length of a country's report entry grows, the AI, SD, or CIRI score is slightly higher. In other words, to a small but statistically significant degree, the human rights score is worse for longer reports, controlling for severity of violation by including the other report score in the regression. Because of the logged transformation of the word count, the coefficient for the logged wordcount is not particularly transparent or easy to interpret.   For heuristic purposes we did the same analysis on the raw wordcount data, and the levels of significance of the coefficients were similar (Table 3, columns 2, 4, 6, and 8). When regressions are run on the raw wordcount data a rough interpretation is possible. Thus, the raw *AI_wordcount* coefficient, with a value of .00008, suggests that for every 100-word extension in length of the AI report, the *AI-pts* score rises (i.e., gets worse) by slightly less than 1/100 of a point on the transformed one-point *AI-pts* scale. A similar, but much smaller, result is reported for the *SD_wordcount* coefficient. Here, the coefficient, at a value of .0000035, suggests that for every 1000-word increase in the length of a country's State Department report entry, the *sd-pts* score rises (i.e., gets worse) by a little over 3/1000 of a point on the transformed one-point scale. Note, however,

that the lengths of the State Department reports do vary by thousands of words, as is apparent in Figure 1.[11]

The only finding suggesting that length is *not* statistically significant at conventional levels is for the test of the association between CIRI's physical integrity scale and the word count of State Department reports. The coefficient for this relationship is not significant and has a large standard error. This lack of relationship suggests that the CIRI approach to scoring may be less responsive to report length than PTS's standards-based scoring, and perhaps that it is less responsive to information effects as we have conceived them in Row 1 of Table 1. However, surprisingly, we find that the CIRI physical integrity score *is* as responsive to logged AI report length as the *AI-pts* score, at a statistically significant level. The coefficient for the raw word count is similar in size to the ones for PTS scores at .00009, suggesting again that for every 100-word extension in length of the AI report, the *CIRI-physint* score rises (i.e., gets worse) by slightly less than 1/100 of a point on the transformed one-point scale. This suggests that to the extent that CIRI also employs intersubjective text-based coding (for example, in judging levels based on the text when numbers are not reported), it also may be vulnerable to this kind of information effect.

To summarize, then, the conclusion we draw from the statistical significance of report length for human rights score is that coding responds not just to intensity of human rights violations but also the amount of reporting on the country. There could be several reasons for an effect like this. The most likely one, we think, is that intersubjective coders, especially those using a standards-based scheme, may simply unwittingly respond in part to length and level of detail as an indication of intensity of violations, even when we control for "actual" violations. Another phenomenon that may be at work is one we call "reputational": that is, if a country has a bad reputation, the report author(s), or coders, of individual reports may intensify their attention to problems in this country year to year. Evidence of this latter effect in the Guatemalan case presented below is suggestive.

---

[11] In fact, the mean word count for State Department entries on Latin America from 1981-2006 is 6078, with a range of between 779 and 37148 words and a standard deviation of 4948. In contrast, the mean word count of all Amnesty International entries on Latin American countries was 1174, with a range from 54 to 3892 words, with a standard deviation of 701.

Our third hypothesis asks whether the scales are appropriately responsive when violations are particularly bad. We are not the first to note that countries falling in the scales' upper levels of repression may actually differ significantly in their actual levels of violence (Hafner-Burton and Ron 2009:381) (Wood and Gibney 2010:13). Cingranelli and Richards noted in an early reference to the CIRI coding strategy that their coding categories are based on 'event-based' criteria, but that the categories they have chosen in using a (0,1,2) scale on the different components of the physical integrity score are more appropriate than a raw count because "these scoring categories fit the accuracy of the data being reported" (Cingranelli and Richards 1999:519). Thus, contrary to accusations that the CIRI codes are too ambitious in presuming to count individual incidents, the rationale appears more modest, an effort to treat the information level of the texts appropriately. However, it is easy to see that a good deal of meaningful variation between countries may be lost in this way. While a standards-based scale with more variation (such as PTS) should be helpful in this regard, PTS may have similar pitfalls at the upper levels of violation (Hafner-Burton and Ron 2009:381). Mainly referring to CIRI, Landman and Carvalho call the problem of lack of variation at one end of scales like this "variance truncation" (2010:90). We note that the artifact of variance truncation results seems to be a kind of stickiness at the upper levels of the scales, making it harder for the scales to register some kinds of improvements.

The Truth Commission data in comparison is revealing on this score. Five Latin American countries – Argentina, Chile, El Salvador, Guatemala, and Peru – have empaneled domestic truth commissions that produced data on deaths and disappearances based on actual lists of names of victims and dates of disappearance or death. All five countries' truth commissions produced definitive works on repression, focused on the most severe forms of human rights violations, and the overall evidence of repression they present is the most reliable available. Chile, Argentina, and Peru had nationally appointed truth commissions; El Salvador and Guatemala had UN-sponsored truth commissions. [12] Other

---

[12] On Argentina: National Commission on Disappeared Persons (CONADEP) (1984). Nunca Más: Informe de la Comisión Nacional Sobre la Desaparición de Personas. Buenos Aires, Editorial Universitaria de Buenos Aires. On Chile: Truth and National Reconciliation Commission (1991). Informe Rettig: Informe de la Comision de Verdad y Reconciliación. Santiago, Chile, Talleres de La Nación.; on

nongovernmental efforts at data collection supplemented the official documentation efforts.[13]   As noted

above, the texts upon which PTS and CIRI are based are contemporaneous to the events they chronicle,

done in the "night and fog" of human rights violations. After the fact, in some cases, it may be possible to

arrive at more precise estimates of human rights violations. This is particularly the case in countries that

have made a transition to democracy and have devoted considerable political will and resources to

documenting past human rights violations, most often through some kind of investigatory commission or

truth commission.[14]   To begin to probe the hypothesis that the data sets are less responsive to higher

violations, we chart the data from the five Latin American truth commissions for comparison with CIRI

and PTS.[15]   (See Figures 2-6.) For the figures, we charted only the AI-PTS score. We have kept the

original scaling of the AI-PTS (1-5) and CIRI (0-8), but inverted the CIRI scores so that, like the PTS and

truth commission data, a higher score indicates worse violations and the lines are easier to compare.

Guatemala: Commission for Historical Clarification (1999). Guatemala: Memory of Silence: Report of the Commission for Historical Clarification. Guatemala City, CEH (Commission for Historical Clarification).. On El Salvador: United Nations (1993). From Madness to Hope: The Twelve War in El Salvador, Report for the Commission for the Truth for El Salvador. 1 April 1993, U.N. Doc. s/25500,1 April 1993. On Peru: Truth and Reconciliation Commission (2003). Informe Final (Final Report). Lima, Peru, Comisión de la Verdad y Reconciliación.

[13] We used one of these reports to compile annual data on Guatemala. The data on disappearances and deaths that we used was compiled in a joint project of the International Center for Human Rights Investigations (*Centro Internacional para Investigaciones en Derechos Humanos* [CIIDH]) and the American Association for the Advancement of Science (AAAS) Ball, P., P. Kobrak, et al. (1999). State Violence in Guatemala, 1960-1996: A Quantitative Reflection. Washington, DC, American Association for the Advancement of Science and Centro Internacional por Investigaciones en Derechos Humanos. CIIDH pooled data from the press and from testimony provided to human rights organizations. This data, because it is broken down by year, can be compared more directly to the PTS and CIRI scales, but the total number of state sponsored deaths and disappearances documented by CIIDH for the period 1959-1995 is 36,906, a level similar to that registered by the U.N. Truth Commission.

[14] General information and links on truth commissions worldwide is available at United States Institute of Peace (USIP) (2010). Truth Commissions Digital Collection. Washington, DC, The Margarita S. Studemeister Digital Collections in International Conflict Management, available at <http://www.usip.org/library/formin.html>.; see also Hayner, P. B. (1994). "Fifteen Truth Commissions--1974 to 1994: A Comparative Study." Human Rights Quarterly **16**(4): 597-655; Hayner, P. B. (2001). Unspeakable Truths: Confronting State Terror and Atrocity. New York, Routledge., and Chapman, A. R. and P. Ball (2001). "The Truth of Truth Commissions: Comparative Lessons from Haiti, South Africa, and Guatemala." Human Rights Quarterly **23**(1): 1-43.

[15] We draw on data from Sikkink, K. (2004). Mixed Signals: U.S. Human Rights Policy and Latin America. Ithaca, NY, Cornell., and add data from Peru.

More and better information on the nature, intensity, and scope of human rights violations is available today for these five cases than in the past. The truth commission reports for these countries probably underestimated the total amount of repression in each country, because they relied on exact lists of names and thus leave out the many victims of repression that never appeared on the lists. The information from these truth commissions confirms some patterns already identified by AI's and the State Department's contemporaneous reports, but in some ways the patterns modify our understandings of the nature of human rights violations in these countries.

Repression took place at somewhat different time periods in the five countries. If we could trace repression on a map of Latin America, it might appear that there were sub-regional waves of repression. The first, a Southern Cone wave, started in Brazil in the late 1960s, moving to Chile and Uruguay by 1973, followed by Argentina and Paraguay after 1976, and ending in Bolivia in 1980-81. The second wave in Central America began in Nicaragua in 1977 and engulfed El Salvador, Guatemala in the late 1970s and early 1980s. The third wave, in the Andean region, came later than the regional trend, beginning in Peru in the mid-1980s and continuing in Colombia and Venezuela, throughout the 1990s. The Andean wave is distinguished from the earlier waves in that repression took place under governments that are formally electoral regimes, if not fully democratic.

While we are not able to discuss each of the truth commission countries here, visual inspection of Figures 2-6 suggests a couple of things. The truth commission data represents large-scale killing and disappearances and, on inspection, the numbers on the graphs are spiked, while the PTS and CIRI codes over the course of the same years are smoother and stay up over more than one time period. The PTS coding is smoothest, without great changes year-to-year. The scores do not change as quickly after the lethal violations diminish. Reputational effects might be the reason for the "stickiness" we observe in comparing the spiky changes in lethality indicated in the truth commission data with the smoother lines created by the PTS and CIRI coding values.

Finally, one of the trickier questions is how and whether the scales' coding methods adequately reflect different kinds of violations. Here we turn to more detailed illustrations for two cases: Brazil and

Guatemala. Amnesty International has always insisted that different kinds of violations are not commensurable, and this is important to keep in mind as we examine these cases and comment on their coding. First, the case of Brazil illustrates how changes in type of repression interact with changes in the reporting capabilities of NGOs over time to produce some surprising equivalencies in coding.   The second case, Guatemala, illustrates the notable and well documented political bias that characterized the U.S. State Department's reports on that country in the 1980s, which made its way into the PTS and CIRI. Like Brazil, the case also demonstrates the changes in domestic NGOs' level of reporting over time, which may contribute to the information effects we can observe.

Brazil, pre- and post-1985

Changes in types of violations become a part of the information effect problem when more information about somewhat less lethal violations may result in equivalent codings. Here, we probe this issue by examining the case of Brazil.   Brazil had an authoritarian military regime from 1964 until 1985. Brazil's 1989 elections were the first to elect a President by direct popular vote since before the military coup, so we might date the complete transition to democracy to that year.   To simplify, we can think of the decade of the 1970s as a period of authoritarian rule, the 1980s as a period of transition towards democracy, called the *abertura*, or "opening," and the decades of the 1990s and 2000s as periods of democratic rule. Since 1995, Brazil has been governed by relatively stable and well regarded governments of the center and the left – led, respectively, by Presidents Fernando Henrique Cardoso (1995-2002) and Luiz Ignacio da Silva (2003-present).

The problem that Brazil poses is that there is a difference between the substantive view of human rights practices in Brazil, and its scores on the human rights measures. Researchers generally believe that human rights practices were worst during the period when General Emílio Médici held power (1968-1974) and improved somewhat during the period of abertura. After the transition to democracy, serious human right problems remained, but most would agree that the democratic period has had better human rights practices than the military regime.   The databases, however, indicate that the human rights practices during the military government were better than those of the democratic period. The PTS data

set begins in 1976, so the worst years of repression are not covered, but Brazil has an average PTS score of 3 for the authoritarian and transition decades of the 1970s and 1980s. For reference, these scores are provided in Figure 7. For the fully democratic period of the 1990s and 2000s, however, Brazil has an average PTS score of 4 in the 1990s and 3.9 in the 2000s. In other words, the human rights situation in Brazil was a full point better during the authoritarian and transition period than it has been during the democratic decades.

The CIRI coding for Brazil shows the same tendency. Again, CIRI begins in 1981, but the average CIRI physical integrity rights score for Brazil for part of the military government period that is covered (1981-1985) is 4.4, the same as the entire 1980s period. The average CIRI score for both the 1990s and the 2000s is 3.3. Because CIRI breaks the score down further by specific type of violation, we can see more precisely the kinds of violations that are driving the scores by examining the different components of the scale. In the case of post-transition Brazil, what is clearly driving the poor physical integrity scores is the use of extrajudicial killings and torture. The score shows that the democratic Brazilian governments rarely practiced disappearances or held political prisoners, a problem during the earlier period, but there has been an unrelenting practice of killing and torture. The CIRI definition of extrajudicial killing explicitly includes "deliberate, illegal, and excessive use of lethal force by police," whether against criminal suspects, detainees, prisoners or others.

How can we explain these differences? First we should clarify that the PTS and CIRI scales are measures of physical integrity, not of civil liberties and democracy, so they disregard regime type when measuring physical integrity violations. There are indeed cases where human rights practices are worse during democratic regimes than authoritarian ones. But we believe that other factors are at stake here that illustrate the information effects we highlight. While Brazil did not have a full-fledged official truth commission, it did have a "Special Commission on Deaths and Disappearances," which prepared a 493-page document listing all known cases of proven deaths and disappearances, before, during, and after the military government (Secretaria Especial dos Direitos Humanos do Presidência da Republica 2007). The report makes clear that the worst period for deaths and disappearances was the period from 1971 to 1974,

that the practice of state killing and disappearing of political opponents was rarely used after 1979, and that there have not been any cases after the 1985 transition to democracy.

Why then do the data seem to reflect that respect for human rights was worse in the 1990s and 2000s than during the military regime? The most plausible answer is that human rights organizations have expanded their focus from a narrow concentration on direct government responsibility for 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 excess use of lethal force, and the duty of the state to prevent, investigate and prosecute violence by non-state actors.

To some degree, the focus of human rights organizations has expanded. In its first specific report on Brazil in 1976, Amnesty focused on gross human rights violations, especially political imprisonment, torture, and summary executions committed directly by state officials, and there were few sources of human rights information either outside of or inside Brazil.   Although Amnesty mentions 800 political prisoners, four deaths of prisoners in captivity, 23 disappeared people, and systematic torture throughout the country, PTS codes this report as a 3. Amnesty asked the authoritarian Brazilian government to make on-site visits but it was not permitted to do so. The State Department's first specific report on Brazil in 1981 was short and largely positive, commending the government on improvements. This focus on improvements of basic rights is consistent with the later Special Commission Report, which shows no state sponsored deaths and disappearances in 1981.

By 1987, however, in the period of transition under the government of Jose Sarney, both Amnesty and the State Department take a harsher tone. Amnesty mentions its specific missions to Brazil and a resulting report on rural killings. Amnesty clarified that such killings were carried out by "hired gunmen in the pay of local landowners," but stressed that it is concerned about the "persistent failure by local and state authorities to investigate these killings effectively or to bring criminal prosecutions, with the result that those responsible acted with impunity and further abuses were encouraged."   A second AI investigation looked into the practice of torture and ill-treatment of detainees in police stations and prisons throughout the country and killing of suspects.

From the point of view of human rights work, this is an encouraging development. AI moved from a quite narrow focus on state sponsored imprisonment, killing, and torture, to a focus on state failure to prevent, investigate, and prosecute violence, in the case of the rural killings, and a focus on police brutality and excessive use of lethal force with criminal suspects. From a point of view of human rights measurement, AI's expansion of its mandate and better capacity to do on-site investigations has led to an increasing documentation of human rights violations that could be misinterpreted if researchers rely only on the coding. Further evidence that this expansion has occurred is provided in the State Department entry on Brazil for 1987, which mentions the importance of NGO activism.

> Although political rights are protected in Brazil today, problems remain in other human rights areas – most notably in the treatment of peasants involved in land disputes, common criminals, and Indians. Common criminals in the custody of police are frequently beaten, and some criminals not in custody are killed anonymously. These phenomena are not new, but their frequency apparently is increasing. These incidents have received considerable public attention in the past year due to the increasing assertiveness of local human rights organizations and the press, as well as growing public concern about criminal violence. (1987 entry on Brazil)(United States. Department of State 1979-)

The PTS codes the 1987 State Department report as a 4. This focus on police killings and torture has characterized the bulk of reports on Brazil throughout the 1990s and 2000s, and organizations within Brazil have turned their attention to such police violence. For example, there is an organization called *Observatório das Violências Policiais-SP* (Observatory of Police Violence in Sao Paulo) that uses news sources to provide a monthly report of all victims of police violence. They argue that all of these deaths are "summary executions," and express the high level of lethal force of the Brazilian police. Many of these killings are of poor and marginal populations living in the slums of Sao Paulo.

This excellent human rights work thus calls our attention to human rights violations against groups not previously the focus on the human rights movement. These are very serious violations of human rights, but the question from the point of view of a researcher sensitive to information effects is "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.

Guatemala

Guatemala also illustrates these kinds of information effects, as well as early political bias muddying State Department reporting. The wave of state-sponsored violence that swept Guatemala from 1980 until 1983 was on an order of magnitude that dwarfed the repression that came before and after this period. Guatemala is the only case in Latin America to which the term "genocide" persuasively applies, not only for the very large number of victims, but also because indigenous people were intentionally targeted by the state.

In 1999, the U.N.-sponsored truth commission in Guatemala, the Historical Clarification Commission, issued a report confirming that the Guatemalan state, in the context of an armed confrontation with guerrilla groups, had carried out genocide against its Mayan population, resulting in as many as 200,000 deaths. In its actual documentation, however, the Commission registered 23,671 victims of arbitrary execution, and 6,159 disappearances, for a total of almost 30,000 deaths and disappearances. The bulk of these violations took place during the relatively short period of the successive military regimes of Romeo Lucas García and General Efraín Ríos Montt from 1980 to 1983. A four volume report published in 1998 by the Human Rights Office of the Roman Catholic Archbishop of Guatemala (ODHAG), *Guatemala: Nunca Más*, estimated that security forces had killed 150,000 people and disappeared 50,000 people during the period from 1960 to 1997 (ODHAG (Oficina de Derechos Humanos del Arzobispado de Guatemala) 1998).

State Department reports for this period, however, dramatically underreported the scale of killing and suggested that it was impossible to know who was responsible for the killings: the government, right wing groups, or the guerrillas. In 1982, the year that truth commission reports retrospectively revealed the highest levels of state sponsored deaths and disappearances ever reported in any country in the hemisphere, the State Department report says:

> With Rios Montt's assumption of power, the situation began to change. The government publicly committed itself to end the abuses of the Lucas Garcia government…. In Guatemalan cities there has been a marked decrease in killings and disappearances since Ríos Montt came to power, although some abuses continue to be reported…. The situation in the countryside, on the other hand, remained unclear for several months as the army fought to regain areas that had been largely abandoned to guerrilla control …. During this period it is believed that members of both the army and the guerrillas were responsible for killings of civilians …. In recent months it has become clearer both that the army has increased its control over areas formerly dominated by the guerrillas, and that its forces have begun to pay greater attention to the importance of proper treatment of the civilian population. Although there continued to be credible reports of human rights abuses by some military units, the overall conduct of the armed forces had improved by late in the year. (entry on Guatemala, 1982) (United States. Department of State 1979-)

There are various possible explanations for the great discrepancy between what the truth commission and ODHAG reports agree on and what the State Department wrote at the time. For some time it has been confirmed that the new Reagan administration's political agenda influenced the report. In its opposition to the Carter Administration's human rights policy, the Reagan administration had decided that the Carter policy had unnecessarily alienated the Guatemalan government: to reestablish a dialogue, the U.S. would need to make some "condition-free demonstrations of our goodwill."[16] Embassies were expected to shape their reporting to the overall policy the administration pursued in the country,[17] and the tone of the State Department's 1982 report on Guatemala appears to reflect that new policy.

The fact that political bias affects the report is clearer when one reads the confidential CIA reports from the embassy at the same period. While the human rights reports express confusion about who is responsible for deaths and disappearances, the CIA cables are straightforward. A CIA cable reported that the Guatemalan army had

> launched a sweep operation into Ixil Province. The commanding officers of the units involved have been instructed to destroy all towns and villages which are cooperating with the EGP [*Ejercito Guerrillo de los Pobres* – the main guerrilla organization in Guatemala at the time] and eliminate all sources of resistance .... When an army patrol meets resistance and takes fire from a town or village, it is assumed that the entire town is

---

[16]United States, Department of State, "Initiative on Guatemala," Cable to Embassy from Latin American Bureau, National Security Archive, April 4 1981.
[17]F. Allen Harris, interviewed by Kathryn Sikkink, March 5, 2003, Washington D.C.

> hostile and it is subsequently destroyed .... The army high command is highly pleased with the initial results of the sweep operation.[18]

The report concluded that "the well documented belief by the Army that the entire Ixil Indian population is pro-EGP has created a situation in which the army can be expected to give no quarter to combatants and non-combatants alike." (Ibid.)

The political bias shows in the PTS score assigned to the text of the State Department report. (Coding values for 1975-1990 are appended at Figure 5.) The State Department PTS score remains at 4 (instead of the worse possible score of 5) throughout the entire period of genocide in Guatemala, from 1980-1983. But after the worst period of killing and disappearance ended, the State Department PTS coding stayed high, at 4 (with an occasional 5) until 1997. So we need to recognize and explain not only the underestimation of human rights violations by the State Department in the period of genocide, but also the "stickiness" of the human rights reporting and coding for 15 years after the genocide. Some of the information effects we discuss here may help explain this feature of the data.

Amnesty International reports from 1978 to 1983 had a content and tone completely different from the State Department reports during the same period. Although AI never uses the term genocide, the reports make clear that AI fully understood the extreme seriousness of the situation in Guatemala at the time. (See also Clark 2001.) AI attributed most human rights violations to state officials or government-sanctioned death squads. Reports highlighted the movement of repression into the rural provinces where indigenous Mayans lived. By 1979, AI says that journalists characterize 1978 as "the most violent year in the country's history" (Amnesty International 1975 -) In 1980, AI writes that there are "no cases of long term political imprisonment," but that "hundreds have been held, interrogated, tortured and eventually killed," and that "an estimated 1,800 people have been disappeared and/or killed." (Amnesty International 1975 -) Reports get worse in 1981 and 1982. When these reports are assigned a PTS code, they receive the worst score, 5, for the deepest period of the genocide (1980-1983). But, the AI reports also use level 5 to code the entire period around that time, 1977-1985, and again in 1990.

---

[18]United States, Central Intelligence Agency, "Guatemalan Army Conducts Sweep Through Ixil Region," National Security Archives, February 1982.

What this pattern suggests is that the coding scale simply cannot encompass a period of genocide or mass murder, nor can it compare such a period to other periods of very high human rights violations. The coding does not respond to extreme worsening or improvement of the human rights situation. Reading AI reports for the late 1970s, it is hard to image how the situation can get worse. When, in the early 1980s, the situation does indeed worsen, as a horrendous situation of human rights violations spirals into genocide, the AI report uses higher estimates of the number of deaths and disappearances, but the AI-PTS coding of the AI report cannot capture the difference.

The CIRI physical integrity score, which begins with the worst score in 1981 for Guatemala, shows a slight improvement in the physical integrity score for 1981 and 1982, the deepest period of the genocide. If a government like Guatemala's in the 1980s is simply killing all its presumed opponents, it may appear as if the physical integrity index improves because there are no political prisoners. The nature of the extreme repression, and possible reflection of political bias, since CIRI relies on the State Department reports as its primary source, may help explain why. In 1982, CIRI gives Guatemala an intermediate score for torture, and in 1983, assigns an intermediate score for political imprisonment, before returning to the worst scores for 1984-1986.

Another probable reason for the stickiness we note for the case of Guatemala is the changing availability of human rights information. During the early 1980s, the repression was so severe that it eliminated or silenced the human rights movement, and thus eliminated the information that would have been generated by this movement. There were no effective local human rights NGOs in Guatemala until the mid-1980s because two of the most prominent (the Guatemalan Commission for the Defense of Human Rights, and the Committee for Justice and Peace) had been forced to close down. In the mid-1980s, human rights organizations began to be able to exist and work again in Guatemala although they still faced profound repression (Americas Watch 1989:43). By the 1990s, a more diverse set of human rights organizations were now working domestically in Guatemala, and their numbers increased. Most importantly, after years of silence, the Roman Catholic Church in Guatemala opened the Archbishop's human rights office referenced above, ODHAG, in 1990. ODHAG became the most prominent and

professional human rights organization in Guatemala and developed and maintained extensive contacts with transnational network actors.[19] Human rights organizations were also able to tentatively "link up" with newly-formed state-level offices, especially the government's Human Rights Ombudsman's office formed as a result of the 1985 constitution. Eventually the Human Rights Ombudsman became a more effective source of human rights information and advocate for human rights, especially from 1989 to 1993, when Ramiro de León Carpio was its head.

The process of redemocratization in Guatemala permitted domestic human rights NGOs to form, led to the creation of state human rights institutions, and in sum created a more information rich environment where both the State Department and AI were well informed about a wide range of ongoing human rights violations. And, once attention is focused on a human rights situation, this attention is often sustained over time, so that a country is scrutinized more carefully for years. NGOs may devote more staff time to that country, and staff develop expertise and contacts that allow them to be particularly well informed about human rights violations in these target or priority countries.

In the case of Guatemala, the political bias of early State Department reports also led to an increased scrutiny of State Department reporting on Guatemala.   For a number of years, major human rights organizations offered annual critiques of the State Department's human rights reports (e.g., Americas Watch, Helsinki Watch et al. 1983). One result was that some Members of Congress began to question the objectivity of State Department reporting: in 1989, Representative John Conyers, chair of a key Congressional subcommittee, requested the General Accounting Office (GAO) to examine the State Department's policies and procedures for preparing the annual Country Reports, including whether the reports were accurate and unbiased, and whether State Department staff had the resources and training they needed to prepare the reports. Guatemala was one of eleven countries the GAO chose to examine in depth. The GAO concluded that the State Department reports on El Salvador, Guatemala, and the Philippines had "excused these governments from responsibility for abuses based on their promises of corrective action" (U.S. General Accounting Office (GAO) 1990:13). In Guatemala, even U.S. embassy

---

[19] Daniel Saxon, interviewed by Kathryn Sikkink, Guatemala City, May 22, 1998.

officials agreed that there had been past bias (Ibid.). The accusations of bias led to greater scrutiny of State Department reports, and greater scrutiny led to more complete reporting. According to the GAO, the 1989 State Department report reflected a more "objective and frank" critical approach that evidenced a break from past policies (Ibid.:15). Another study by a State Department insider said that by the mid-1980s, the team within the State Department that prepared the Country Reports was "increasingly willing and able to resist pressures to slant the Reports for political purposes," thus confirming that in the early 1980s, it did not resist such pressures (Innes de Neufville 1986:690).

To summarize, we argue that an increase of domestic human rights NGOs and increased attention to and scrutiny of State Department reporting led to the production of more information about human rights abuses in Guatemala in the wake of the genocide of the early 1980s. The limitations on both coding schemes' ability to respond to extreme worsening or improvement, as well as the increased information, sustained high levels of coding of human rights violations in Guatemala well into the mid 1990s we have described above. From a human rights point of view, this increased scrutiny and increased information is a positive development. From the point of view of measurement, it suggests some systematic error with sources in both the coding schemes and the texts, particularly after especially repressive periods.

Conclusion

We emphasize that we do not agree with data "skeptics" who would reject the use of human rights measures like PTS and CIRI. Nor do we think it automatic that, given the information effects we document, case studies will always be definitive. We believe it is unproductive to characterize the issues we raise as a conflict between qualitative and quantitative approaches. We are interested in advancing more systematic and precise arguments about how and why the data may be biased or subject to measurement error, and how such problems might be addressed. The present study therefore has investigated the nature and limitations of the source material itself, as well as the two methods of coding. We investigated variation in the reports over time and compare them with data from other sources, focusing on the Latin American region. We also analyzed differences between the PTS approach and the

CIRI approach to coding the source material. We conclude by summarizing our findings and we suggest here some ways scholars can be careful users of these data.

What can the responsible researcher do? Scholars need to be thoughtful and informed users of the CIRI and PTS data and to realize that they are based on contemporaneous documents. The numbers are potentially useful generalizations but should not be treated as uniquely objective measures of human rights. Scholars should be aware of the kinds of problems with the data that we spell out, and thus the strengths and weaknesses of the measures. For example, we believe that the measures are not well suited to examine extreme episodes of human rights violations like genocide and mass killing or relative "improvements" that happen after such extreme episodes. The measures are unable to encompass such extreme episodes, and tend towards stickiness after such episodes. The stickiness is something that deserves further analysis than we have been able to accomplish, but we would note that in our visual comparison with truth commission data, the PTS measures appeared somewhat better than the CIRI physical integrity scale at noting periods of extreme repression, but the PTS measure also suffered from stickiness and did not seem to change as rapidly. Both scales had similar responsiveness to report length, and we argue that these are indicators consistent with information effects.

Second, the early years of the measures are particularly problematic, because the reports were shortest at this time, fewer human rights NGOs existed to produce good source information, and the political bias in the State Department reports was the strongest. In the case of Guatemala, the contemporaneous Amnesty International reports were much closer to the retrospective truth commission reports than were those of the State Department. If possible, other measures should also be employed. Some scholars, for example, have collected data on other observable implications of respect for specific human rights for their numeric studies (Simmons study (2009) is a good example). Any new data collection efforts should keep in mind and perhaps estimate the likelihood of whether the pitfalls of informational variation may be at play in particular cases.

Third, it is unhelpful and unnecessary to make blanket judgments on quantitative vs. qualitative approaches. It is more important to think about characteristics of the source materials and how they are

used. Case studies helped immensely in early theory building on the role of human rights change, and both approaches can be used to further advance causal understanding. If there are some information effects, then we need to understand them and begin addressing them using the appropriate means. It is not an either-or proposition and we think it is counterproductive to characterize it in this way. As scholars, practioners, and policy makers ask important questions, we need cooperation on how best to make sense of the increasing information that we do have.

In general, we think that the human rights measures should be used with great care when trying to make the point that a human rights situation has not improved. As we have seen, the CIRI measures of individual components are designed in such a way that they would not distinguish any improvements until the number of violations falls below the "frequent" threshold. This rules out a range of improvement that might be very important to the human rights researcher. Similarly, both scales seem to stay high after periods of high levels of abuse. As the comparison with truth commission reports data suggests, the stickiness may extend beyond variance truncation to, possibly, reputational effects. A further line of research would be to develop and test a theory of how this happens in the coding process. Alternately, researchers should be on the lookout for better or companion sources of information when using such indices.

We argue that the information effects that we discuss usually mean that the existence of more information about human rights violations may make it appear that the human rights situation is worse or has not improved from previous years, when we may simply know much more about human rights violations. This is not a problem limited to human rights research, but plagues many other areas of research as well, such as questions about crime rates or the prevalence of certain diseases. Human rights researchers, however, often ignore the relevance of that question to the reliability of empirical findings in our field. Our main purpose then is to raise the issue, begin an analysis, and urge that the issue be seriously considered by all researchers in the field. We believe there are many other creative ways to test the hypotheses we propose here, and we encourage other researchers to do so.

Table 1. Information Effects and Data Set Differences

Type of bias:	Is problem with source text or with measurement?	PTS associated standards-based problem		CIRI associated standards-based or count-based errors	
		expected	**found?**	expected	**found?**
1. *amount* of information included in report varies	source	standards-based coding can mitigate without false precision[*rely on intersubjective coding]	some artifact of length for AI and SD measures	Use of some count-related coding rules may render CIRI less vulnerable to report length, but will still be vulnerable when counts unavailable	some artifact of length, but only to AI report length
2. *type* of information or human rights issues included in report varies	source	standards-based coding of violence may mitigate variation in level of detail without false precision	Brazil example: change in type. Guatemala example: political bias in State Dept. reporting in early years	If souce info is inconsistent across countries or over time, more vulnerable to false precision	Brazil example: change in type. Guatemala example: political bias in State Dept. reporting in early years
3. same coding can apply to *different types of violation or issues*	measurement	measure may not respond to change in context	Brazil example: scores similar although change in types of violations reported	not expected, since coding is subdivided by type	Brazil example: scores similar although change in types of violations reported
4. coding doesn't respond to *extreme worsening or improvement*	measurement	characterization may persist by coder or source; coding may not be equivalent across countries	Truth Comm comparison: responds better to acute violations, but is sticky	may not reflect change in high abuse (variance truncation); coding may not be equivalent across countries	Truth Comm comparison: responds less well to extreme violations
5. *quality* of source information varies in its focus or is biased	source	coding may not be equivalent across countries; mitigated by use of AI and SD	Guatemala example: political bias in State Dept. reporting in early years	count may be vulnerable to false precision or bias	Physical integrity index less responsive to extreme repression of single type

Table 2. Yearly trend in length (word count): estimates for AI and State Department reports

	AI word count	SD word count
Latin America	no signif. trend, (Sample: 1975-2006)	343.87 ($p<.01$) (Sample: 1981-2006)
All countries	-16.80 ($p<.01$) (Sample: 1999-2006)	297.43 ($p<.01$) (Sample: 1999-2006)

Table 3. Effects of report length on PTS and CIRI coding, controlling for human rights violations: Latin America[20]

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	AI-pts (1975-2006)	AI-pts (1975-2006)	SD-pts (1981-2006)	SD-pts (1981-2006)	CIRI-physint (1981-2006)	CIRI-physint (1981-2006)	CIRI-physint (1981-2006)	CIRI-physint (1981-2006)
AI-pts$_{t-1}$	**.3494 (.0249)	**.3357 (.0416)						
SD-pts$_{t-1}$			**.3764 (.0504)	**.3927 (.0494)				
CIRI-physint$_{t-1}$					**.3437 (.0702)	**.3439 (.0701)	**.3193 (.0556)	**.3088 (.0521)
AI word count, logged	**.0820 (.0159)		**.0437 (.0121)				**.0835 (.0182)	
AI word count, raw		**.00008 (.000018)						**.00009 (.00002)
SD word count, logged					.0040 (.0133)			
SD word count, raw				**.0000035 (.0000013)		-.00000036 (.0000012)		
AI_pts			**.3499 (.0458)	**.3458 (.0458)	**.3450 (.0607)	**.3429 (.0605)		
SD_pts	**.3636 (.0473)	**.3547 (.0456)					**.3123 (.0408)	**.3012 (.0415)
Constant	**-.4165 (.1130)	**.0624 (.0279)	**-.2780 (.1012)	**.0717 (.0204)	**.1070 (.1214)	**.1461 (.0340)	**-.4071 (.1251)	**.0732 (.0308)
	R^2 within: .48 R^2 between: .95 R^2 overall: .72	R^2 within: .48 R^2 between: .96 R^2 overall: .72	R^2 within: .49 R^2 between: .98 R^2 overall: .78	R^2 within: .36 R^2 between: .94 R^2 overall: .69	R^2 within: .36 R^2 between: .95 R^2 overall: .66	R^2 within: .36 R^2 between: .95 R^2 overall: .66	R^2 within: .39 R^2 between: .92 R^2 overall: .68	R^2 within: .39 R^2 between: .94 R^2 overall: .69
	observations: 641 groups: 30	observations: 641 groups: 30	observations: 597 groups: 30	Observations: 597 groups: 30	observations: 515 groups: 25	observations: 515 groups: 25	observations: 545 groups: 26	observations: 545 groups: 26

** = significant at *p*<.01; *=significant at *p*<.05

---

[20] Results based on pooled cross-national time series regression with fixed effects and robust standard errors.

Figure 1. Graph of mean word counts for Latin America and global sample (the global sample mean includes the Latin American countries)
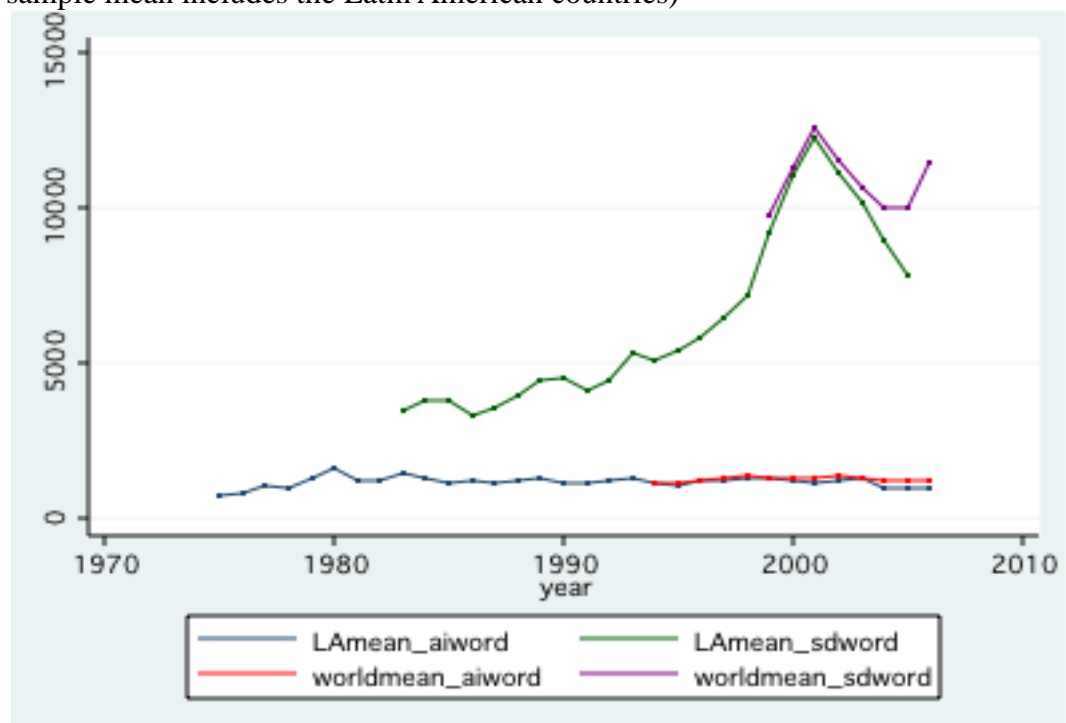
Figure 2. Chile.
(Note: CIRI score represents the added disappearance and killing scores (maximum: 4), reversed
to show 0 as best score, to match direction of PTS and Truth Commission Data.)

Chile: Truth Commission Record of Deaths and Disappearances Compared to PTS and CIRI

Figure 3. Argentina
(Note: CIRI score represents the added disappearance and killing scores (maximum: 4), reversed to show 0 as best score, to match direction of PTS and Truth Commission Data.)
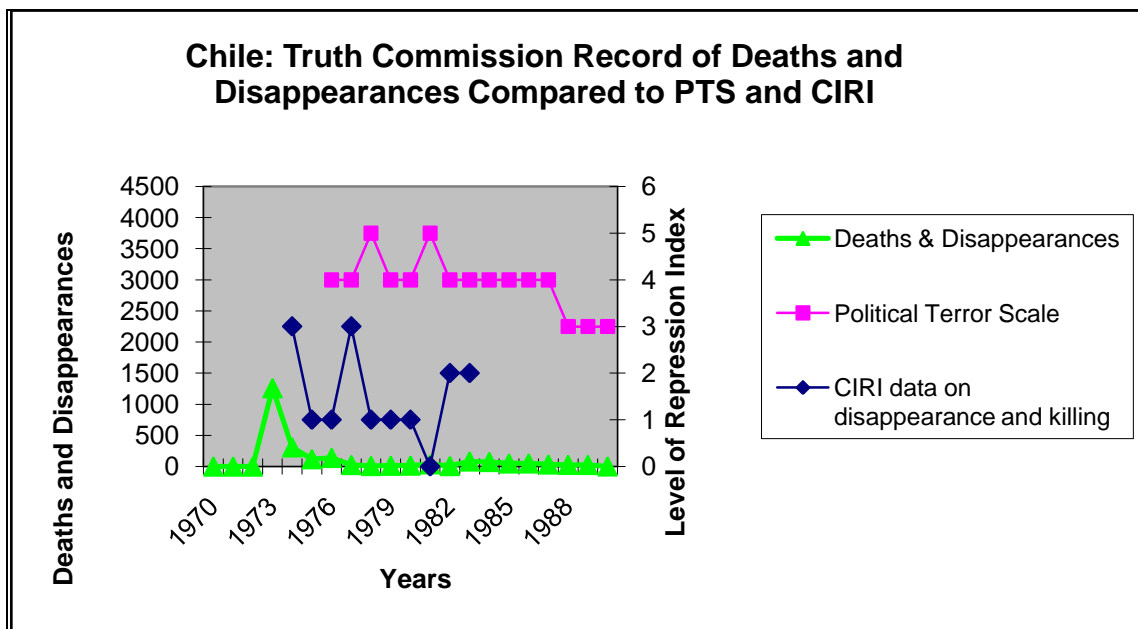

Argentina: Truth Commission Record of Death and Disappearances compared to PTS and CIRI indexes

Legend:
- Number of Disappearances
- Political Terror Scale (based on AI reports)
- CIRI data on disappearance and killing

Figure 4. El Salvador

(Note: CIRI score represents the added disappearance and killing scores (maximum: 4), reversed to show 0 as best score, to match direction of PTS and Truth Commission Data.)
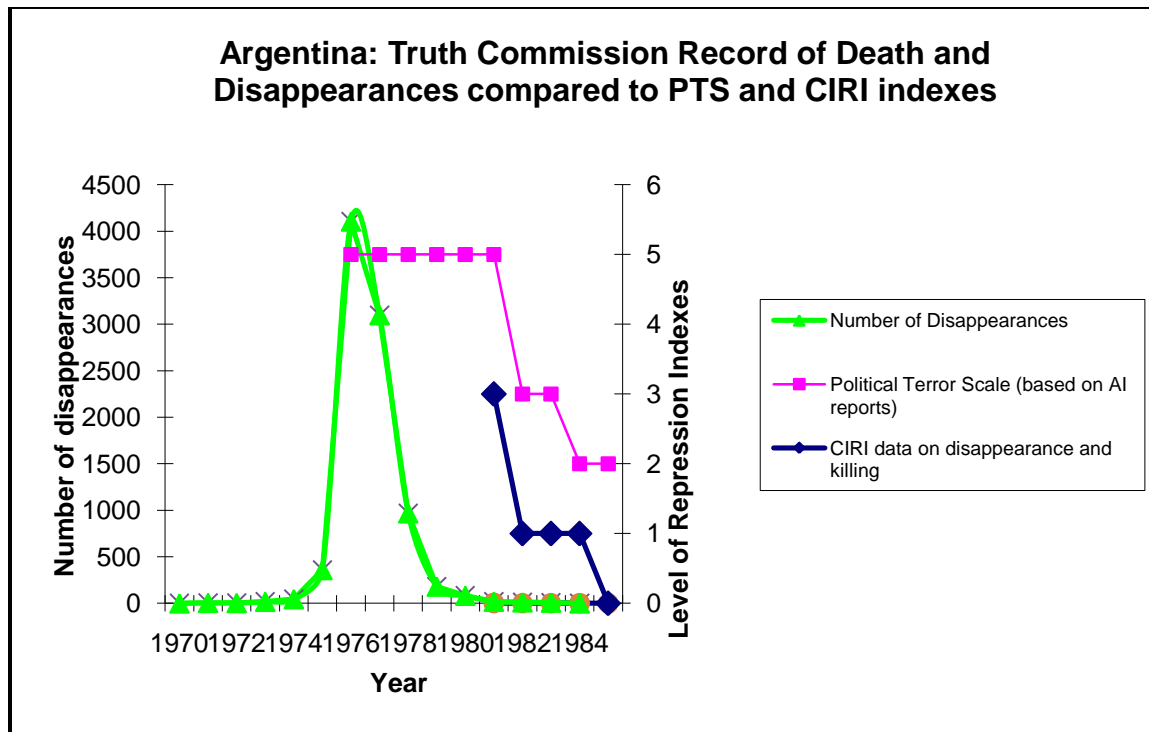
Figure 5. Guatemala, raw data and graph
(Note: CIRI score represents the added disappearance and killing scores (maximum: 4), reversed to show 0 as best score, to match direction of PTS and Truth Commission Data.)
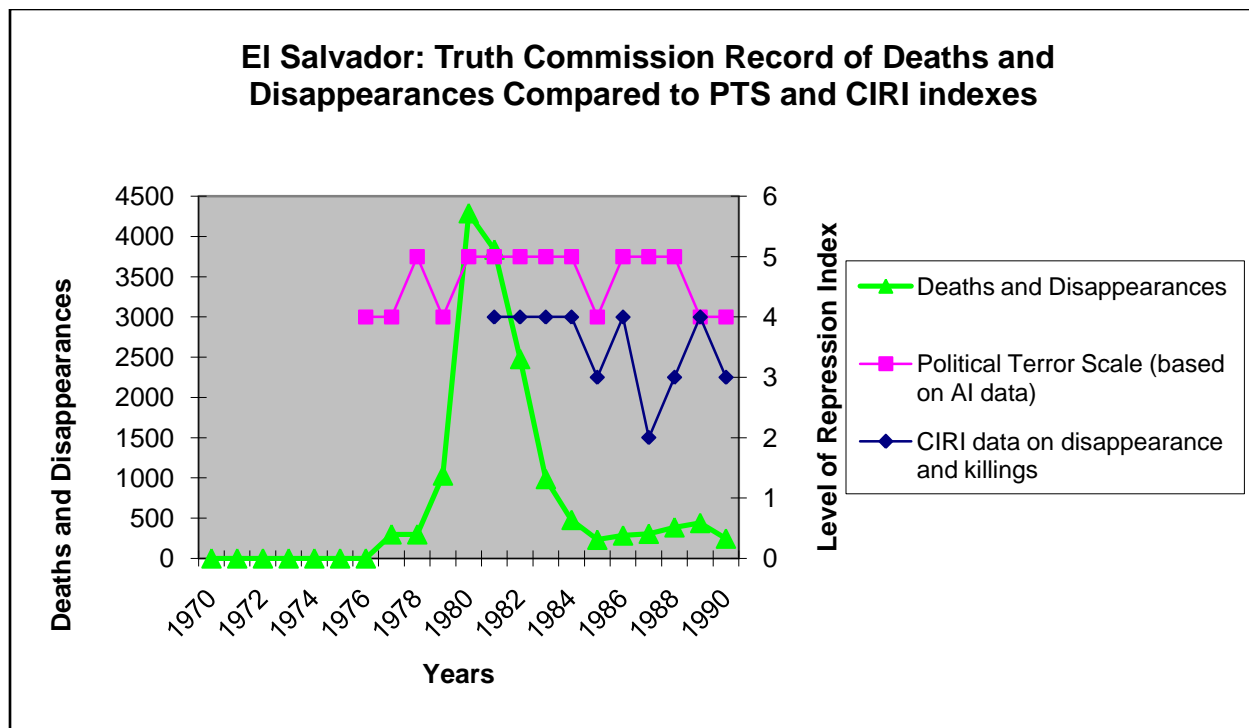
Guatemala, 1976-1995

Year	Truth Commission Count of Deaths and Disappearances	PTS (1=best; 5=worst)		CIRI	
		AI – PTS	SD-PTS	Physical Integrity Scale (inverted: 8=worst; 0=best)	Disappearance and Political Killing (inverted: 4=worst; 0=best)
1970	301				
1971	410				
1972	355				
1973	286				
1974	139				
1975	63				
1976	173	4			
1977	227	5			
1978	203	5			
1979	181	5			
1980	2,349	5	4		
1981	3,736	5	4	8	4
1982	17,953	5	4	7	4
1983	1,890	5	4	7	4
1984	869	5	4	8	4
1985	501	5	4	8	4
1986	296	4	4	8	4
1987	392	4	4	4	2
1988	352	4	4	7	3
1989	409	4	4	6	4
1990	598	5	5	7	4
1991	354	4	4	6	4
1992	326	4	4	6	3
1993	160	4	4	4	3
1994	250	4	5	6	4
1995	95	4	5	6	4

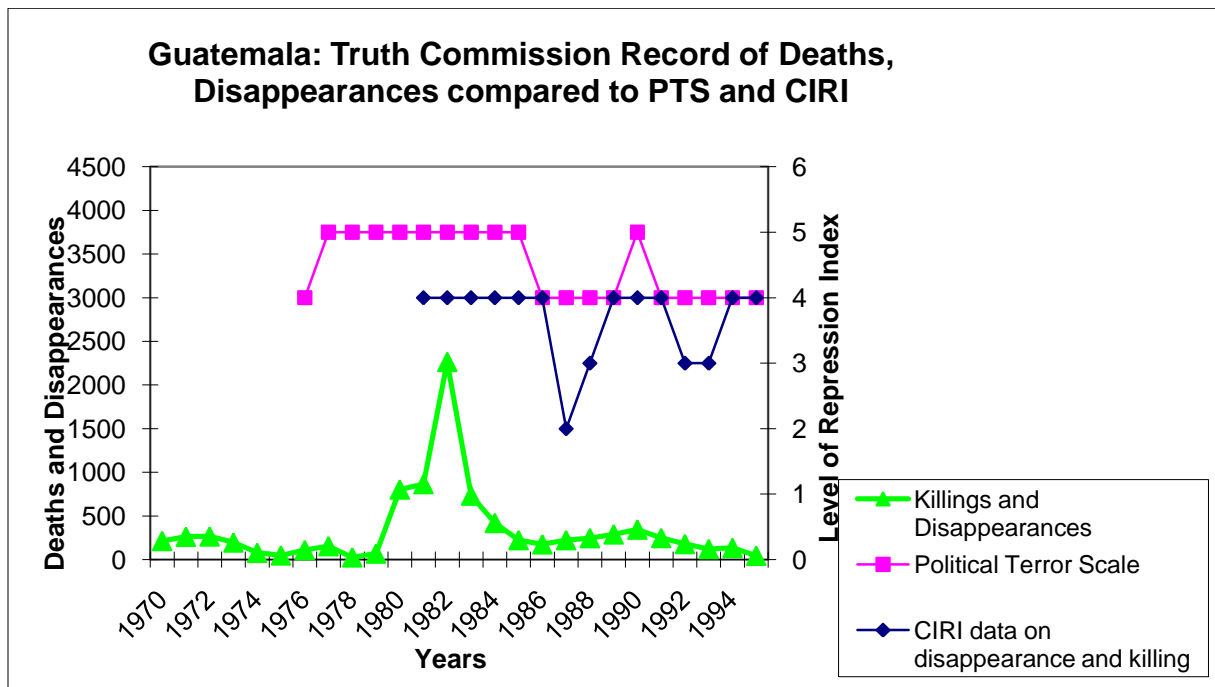(PTS data on graph corresponds to AI-PTS.)

Figure 6. Peru
(Note: CIRI score represents the added disappearance and killing scores (maximum: 4), reversed
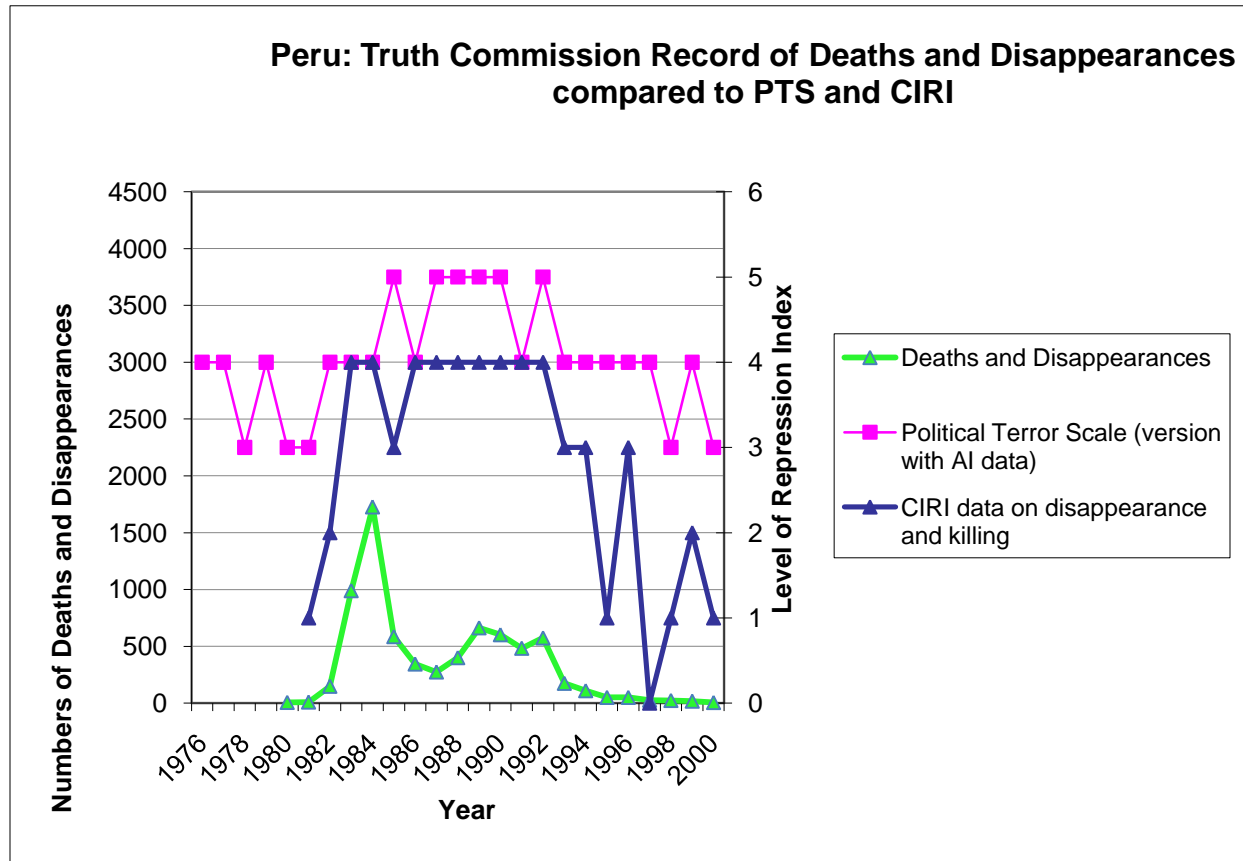to show 0 as best score, to match direction of PTS and Truth Commission Data.)

Peru: Truth Commission Record of Deaths and Disappearances compared to PTS and CIRI

Figure 6. Brazil Coding, 1976-2006.
(Shown is original coding. PTS: 5= worst; 1=best; CIRI physical integrity: 0=worst, 8=best; other CIRI scales: 0=worst, 2=best)

Brazil, 1976-2006

	AI	SD	CIRI Physical Integrity Index	CIRI Disappearances	CIRI Political Killings	CIRI Political Imprisonment	CIRI Torture
1976	3	2					
1977	4	3					
1978	4						
1979	3	2					
1980	3	2					
1981	2	2	4	2	0	1	1
1982	3	2	4	2	0	2	0
1983	4	2	5	2	1	2	0
1984	4	2	4	2	0	2	0
1985	4	3	5	2	1	2	0
1986	3	3	5	2	1	2	0
1987	4	4	4	2	0	1	1
1988	3	3	5	2	1	2	0
1989	4	3	4	2	0	2	0
1990	5	4	3	1	0	2	0
1991	5	4	4	2	0	2	0
1992	5	4	3	1	0	2	0
1993	4	4	3	1	0	2	0
1994	4	4	4	2	0	2	0
1995	4	4	2	1	0	1	0
1996	4	4	4	2	0	2	0
1997	4	3	4	2	0	2	0
1998	3	3	2	1	0	1	0
1999	4	4	4	2	0	2	0
2000	4	3	2	2	0	0	0
2001	4	4	3	2	0	1	0
2002	4	4	4	2	0	2	0
2003	4	4	4	2	0	2	0
2004	4	4	4	2	0	2	0
2005	4	4	4	2	0	2	0
2006	4	4	2	1	0	1	0

References

Americas Watch (1989). Persecuting Human Rights Monitors: The CERJ in Guatemala. Washington, DC, Americas Watch.

Americas Watch, Helsinki Watch, et al. (1983). A Critique of the Department of State's country reports on human rights practices for 1982. New York, Americas Watch, Helsinki Watch, and Lawyers Committee for International Human Rights.

Amnesty International (1975 -). Amnesty International Annual Report. London, Amnesty International.

Baehr, P. R. (1994). "Amnesty International and Its Self-Imposed Limited Mandate." Netherlands Quarterly of Human Rights 12(1): 5-21.

Ball, P., P. Kobrak, et al. (1999). State Violence in Guatemala, 1960-1996: A Quantitative Reflection. Washington, DC, American Association for the Advancement of Science and Centro Internacional por Investigaciones en Derechos Humanos.

Chapman, A. R. and P. Ball (2001). "The Truth of Truth Commissions: Comparative Lessons from Haiti, South Africa, and Guatemala." Human Rights Quarterly 23(1): 1-43.

Cingranelli, D. and D. L. Richards (1999). "Respect for Human Rights after the End of the Cold War." Journal of Peace Research 36(5): 511-534.

Cingranelli, D. and D. L. Richards. (2008). "The Cingranelli-Richards (CIRI) Human Rights Dataset, http://www.humanrightsdata.org." Retrieved 21 February, 2008, from http://www.humanrightsdata.org.

Cingranelli, D. L. and D. L. Richards (2010). "The Cingranelli and Richards (CIRI) Human Rights Data Project." Human Rights Quarterly 32(2): 401-424.

Clark, A. M. (2001). 'A Calendar of Abuses': Amnesty International's Campaign onGuatemala. NGOs and Human Rights: Promise and Performance. C. E. Welch, Jr. Philadelphia, University of Pennsylvania Press: 55-70.

Clark, A. M. (2001). Diplomacy of Conscience: Amnesty International and Changing International Human Rights Norms. Princeton, NJ, Princeton University Press.

Commission for Historical Clarification (1999). Guatemala: Memory of Silence: Report of the Commission for Historical Clarification. Guatemala City, CEH (Commission for Historical Clarification).

Dezalay, Y. and B. Garth (2006). "From the Cold War to Kosovo: the Rise and Renewal of the Field of International Human Rights." Annual Review of Law and Social Science 2: 231-255.

Gibney, M., L. Cornett, et al. (2008). "Political Terror Scale, 1980-2006." Retrieved 24 February, 2008, from http://www.politicalterrorscale.org.

Goodman, R. and D. Jinks (2003). "Measuring the Effects of Human Rights Treaties." European Journal of International Law 14(1): 171-183.

Griesgraber, J. M. (1983). Implementation by the Carter Administration of Human Rights Legislation Affecting Latin America. Washington, D.C., Georgetown University, Ph.D. Dissertation.

Hafner-Burton, E. M. and J. Ron (2009). "Seeing Double: Human Rights Impact through Qualitative and Quantitative Eyes." World Politics 61(2): 360-401.

Hafner-Burton, E. M. and K. Tsutsui (2005). "Human Rights in a Globalizing World: The Paradox of Empty Promises." American Journal of Sociology 110(5): 1373-1411.

Hathaway, O. A. (2002). "Do human rights treaties make a difference?" Yale Law Journal 111(8): 1935-2040.

Hayner, P. B. (2001). Unspeakable Truths: Confronting State Terror and Atrocity. New York, Routledge.

Hopgood, S. (2006). Keepers of the Flame: Understanding Amnesty International. Ithaca, NY, Cornell.

Human Rights Internet Reporter (1990). Human Rights Directory: Latin America and the Caribbean, HRI. 13.

Innes de Neufville, J. (1986). "Human Rights Reporting as a Policy Tool: An Examination of the State Department Country Reports." Human Rights Quarterly 8(4): 681-699.

Keck, M. and K. Sikkink (1998). Activists beyond Borders. Ithaca, NY, Cornell University Press.

Keith, L. C. (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.

Korey, W. (1998). NGOs and the Universal Declaration of Human Rights : a curious grapevine. New York, St. Martin's Press.

Landman, T. (2005). Protecting human rights : a comparative study. Washington, D.C., Georgetown University Press.

Landman, T. and E. Carvalho (2010). Measuring Human Rights. New York, Routledge.

Landman, T. and J. Häusermann (2003). Map Making and Analysis of the Main International Initiatives on Developing Indicators on Democracy and Good Governance Colchester, U.K., University of Essex, Human Rights Centre**:** 92pp.

National Commission on Disappeared Persons (CONADEP) (1984). Nunca Más: Informe de la Comisión Nacional Sobre la Desaparición de Personas. Buenos Aires, Editorial Universitaria de Buenos Aires.

Neumayer, E. (2005). "Do international human rights treaties improve respect for human rights?" Journal of Conflict Resolution **49**(6): 925-953.

ODHAG (Oficina de Derechos Humanos del Arzobispado de Guatemala) (1998). Guatemala Nunca Más. Guatemala City, ODHAG, Proyecto Interdiocesano de Recuperación de la Memoria Histórica.

Poe, S. C., S. C. Carey, et al. (2001). "How are These Pictures Different? A Quantitative Comparison of the US State Department and Amnesty International Human Rights Reports, 1976-1995." Human Rights Quarterly **23**(3): 650-677.

Poe, S. C., C. N. Tate, et al. (1999). "Repression of the human right to personal integrity revisited: A global cross-national study covering the years 1976-1993." International Studies Quarterly **43**: 291-313.

Ron, J., H. Ramos, et al. (2005). "Transnational Information Politics: NGO Human Rights Reporting, 1986-2000." International Studies Quarterly **49**(3): 557.

Secretaria Especial dos Direitos Humanos do Presidência da Republica (2007). Direto à Memoria e à Verdade: Comissão Especial sobre Mortos e Desaparecidos Políticos. Brasilia.

Sikkink, K. (2004). Mixed Signals: U.S. Human Rights Policy and Latin America. Ithaca, NY, Cornell.

Simmons, B. (2009). Mobilizing for Human Rights: International Law in Domestic Politics. Cambridge, U.K., Cambridge University Press.

Truth and National Reconciliation Commission (1991). Informe Rettig: Informe de la Comision de Verdad y Reconciliación. Santiago, Chile, Talleres de La Nación.

Truth and Reconciliation Commission (2003). Informe Final (Final Report). Lima, Peru, Comisión de la Verdad y Reconciliación.

U.S. General Accounting Office (GAO) (1990). Human Rights: State Department's Commitment to Accurate Reporting Has Increased. Washington, DC, Report to the Chairman, Legislation and National Security Subcommittee, Committee on Government Operations, House of Representatives, September 1990.

United Nations (1993). From Madness to Hope: The Twelve War in El Salvador, Report for the Commission for the Truth for El Salvador. 1 April 1993, U.N. Doc. s/25500,1 April 1993.

United States Institute of Peace (USIP) (2010). Truth Commissions Digital Collection. Washington, DC, The Margarita S. Studemeister Digital Collections in International Conflict Management, available at <http://www.usip.org/library/formin.html>.

United States. Department of State (1979-). Country reports on human rights practices. Country reports on human rights practices : report submitted to the Committee on Foreign Affairs, U.S. House of Representatives and Committee on Foreign Relations, U.S. Senate by the Department of State in accordance with sections 116(d) and 502B(b) of the Foreign Assistance Act of 1961, as amended. Washington DC, U.S. GPO.

Weissbrodt, D. (2003). "Do Human Rights Treaties Make Things Worse?" Foreign Policy **134**(January/February): 88-89.

Wiseberg, L. S. and H. M. Scoble, Eds. (1981). <u>Human Rights Directory: Latin America, Africa, and Asia</u>. Washington, DC, Human Rights Internet.

Wood, R. and M. Gibney (2010). "The Political Terror Scale (PTS): A Re-introduction and a Comparison to CIRI." <u>Human Rights Quarterly</u> **32**(2): 367-400.