Supplementary Materials for "Law and Politics in the Inter-American Court of Human Rights. A New Database on Judicial Behavior and Compliance in the IACtHR"

Øyvind Stiansen   Daniel Naurin   Live Standal Bøyum

May 10, 2020

A   Codebook

This codebook describes the database on *Judicial Behavior and Compliance in the Inter American Court of Human Rights (IACtHR)*. The database includes information about IACtHR judgments that have reached the stage of a merits judgments and contains information about the procedure before the Court, alleged violations of the American Convention of Human Rights, judicial votes, separate opinions, the outcomes of merits decisions, remedial orders, compliance, information about victims, preliminary objections filed by the respondent state, and *amicus curiae* briefs.

The main data source is the detailed and systematic case summaries of the IACtHR's judgments that the Loyola Law School, Los Angeles team, under the leadership of Prof. Cesare Romano, has produced (https://iachr.lls.edu/database). These reports include information about the cases, brief summaries of the arguments made by the Inter American Commission of Human Rights (the Commission) and the IACtHR, preliminary objections, amicus curiae briefs, votes and dissenting opinions, remedies decided and evaluations made on compliance. The case summaries cover cases from 1988 (first judgment

of the Court) to 2016. To date, The Loyola Law School, Los Angeles has produced 201 case summaries to date, which means that this is also the number of cases in our database. The database is currently complete for all cases that received their merits judgment by the end of 2013, but must for later cases be considered a convenience sample.

In some instances we have also supplemented the case summaries from Loyola Law School, Los Angeles with information from the Court's original documents on the judgments that are available from the Court's own web page (`http://www.corteidh.or.cr/index.php/en`). For example, when a judge dissents from the majority of the judges, we have consulted the original documents to find out what part of the judgment the judge dissented against.

The database is organized into 9 different tables. Below follows explanation of the variables and specific coding instructions for each of the tables.

Cases

The following variables are measured on the case level and contain basic information about the case, its procedure before the Court, and the judges participating in the different stages of the proceedings. The cases are the units of analysis in this part of the database. The variable CaseID can be used to merge this table with all other tables in the database. The Correlates of War country code for the respondent state is included to allow users to merge our data with other dataset containing information at the country or country-date level.

Table A1: Codebook for Cases table

Variable name	Variable description	Coding instructions
CaseID	The ID variable for the cases. The units of analysis in this part of the data.	Automatically generated ID variable
CaseTitle	The title of the case	The title as it is written in the official judgments
RespondentState	The state that the Commission has alleged violations against, and that is the respondent in the case.	The name of the respondent state as a string variable (chosen from a drop-down menu).
COWcode	The Correlates of War country code for the respondent state.	Numeric country code (automatically generated)
DatePetition	The date the petition is submitted to the Commission	dd.mm.yyyy
DateSubmission	The date when the Commission submitted the case to the Court	dd.mm.yyyy
DateMeritsRuling	The date the Court issued its judgment on the Merits	dd.mm.yyyy
JudgesInMeritsHearing	The judges participating in the Judgment on the merits	Vector of judge names chosen from a drop-down menu of all judges that have been members of the Court.
AdhocJudgesInMeritHearing	Is there an *ad hoc* judge participating in the judgment?	The name of the *ad hoc* judge, or "Not relevant" if there is no *ad hoc* judge in the judgment. Chosen from a drop-down menu
RecusingJudgeInMeritsHearing	A judge might recuse himself/herself from the judgment. This is most usually the case if a judge's nationality is the same as the respondent state	Judge name chosen from a drop-down menu of all judges that have been members of the Court or "Not relevant" if there are no recusing judges in the judgment.
AbstainingJudgeInMeritsHearing	A judge might abstain from participating in the judgment because of "reasons beyond his/her control"	Vector of judge names chosen from a drop-down menu of all judges that have been members of the Court or "Not relevant" if there are no abstaining judges in the judgment.

iii

iv

DateRemediesHearing	The date the Court issued its judgment on Reparations and Costs. The judgment on the merits and the judgment on reparation and cost may not be on the same date. This is especially the case for the Court's early judgments. This is not always clear in the case summaries. If unclear, we consulted the IACtHR judgments.	dd.mm.yyyy
JudgesInRemediesHearing	The judges participating in the Judgment on the Reparation and Costs.	Vector of judge names chosen from a drop-down menu of all judges that have been members of the Court.
AdhocJudgesInRemediesHearing	Is there an *ad hoc* judge participating in the judgment?	The name of the *ad hoc* judge, or "Not relevant" if there is no *ad hoc* judge in the judgment. Chosen from a drop-down menu
RecusingJudgesInRemediesHearing	A judge might recuse himself/herself from the judgment. This is usually the case if a judge is a national of the respondent State	Vector of judge names chosen from a drop-down menu of all judges that have been members of the Court or "Not relevant" if there are no recusing judges in the judgment.
AbstainingJudgesInRemediesHearing	A judge might abstain from participating in the judgment because of "reasons beyond his/her control"	Vector of judge names chosen from a drop-down menu of all judges that have been members of the Court or "Not relevant" if there are no abstaining judges in the judgment.
InterpretationRequest	Did the respondent state, the Commission and/or the representatives of the victims ask for an interpretation of the Court's judgment(s)?	Select one or more of the following options: "By Respondent State" "By Victims' representatives" "By Commission" "No"

InterpretationGranted	The Court can rule the interpretation request inadmissible or admissible. We use the alternative "not relevant/no request" when there has been no request for interpretation of the judgment	Select one of the following options: "Admissible" "Not admissible" "Not relevant/no request"
InterpretationSubject	What was the subject of the interpretation request? To categorize the subject of the interpretation request, we read the section "Interpretation and Revision of the Judgment" in the Loyola Law School, Los Angeles case summaries.	Select one or more of the following options: "Composition of court" "Assessment of evidence" "Preliminary objections" "Merits decision" "Remedy decision" "Compliance decision" "Other" "Not relevant"
JudgesInInterpretationHearing	The judges participating in the interpretation hearing. It is not always clear from the Loyola Law School, Los Angeles Case summaries who are the participating judges in the interpretation hearing. We therefore consulted the original Court documents to find this information.	Vector of judge names chosen from a drop-down menu of all judges that have been members of the Court.
AdhocJudgesInInterpretationHearing	Was there an *ad hoc* judge in the interpretation hearing?	The name of the *ad hoc* judge, or "Not relevant" if there is no *ad hoc* judge in the judgment. Chosen from a drop-down menu
RecusingJudgesInInterpretationHearing	A judge might recuse himself/herself from the interpretation hearing. This is most usually the case if the judge's nationality is the same as the respondent state	Vector of judge names chosen from a drop-down menu of all judges that have been members of the Court or "Not relevant" if there are no recusing judges in the judgment.

v

AbstainingJudgesInInterpretationHearing	A judge might abstain from participating in the judgment because of "reasons beyond his/her control"	Vector of judge names chosen from a drop-down menu of all judges that have been members of the Court or "Not relevant" if there are no abstaining judges in the judgment.
AcceptInternationalResponsibility	Did the State accept international responsibility?	Select one of the following option: "Yes" "Partly" "No" "Unclear"

Victims

The following variables are measured at the level of the alleged victim. The VictimID variable allows merging this table with the table containing data on merits decisions. CaseID allows merging with other tables in the database.

Table A2: Codebook for Victims table

Variable name	Variable description	Coding instructions
VictimID	The units of the analysis in this part of the data	Automatically generated ID variable
CaseID	The case that the victim is part of	ID variable from the Case table
VictimName	Name of the victim(s) in the case.	Name of the victim(s) in the case as a character string.
CollectiveVictim	When the victims are a group, and the individuals are not identified, we code these groups as collective victims. For example, when the victims are a group of indigenous people, an organization, or when the victims are a family. Often, the next of kin of the victims are listed as victims in the judgment. In these cases, we code the next of kin collectively and code them as a collective victim.	Select one of the following options: "Yes" "No" "Unclear"
NaturalPerson	Yes, if the victim is a person, no if the victim is an organization.	Select one of the following options: "Yes" "No" "Unclear"
Female	Is the victim male or female? When we have coded the victims as collective victims, we always code this variable as "unclear".	Select one of the following options: "Male" "Female" "Unclear"
CitizenOfRespondentState	Is the victim a citizen of the respondent state? Often, the nationalities of the victims are not specified. In these cases, we code "No info".	Select one of the following options: "Yes" "No" "No info"
RepresentedBy	Who represented the victims?	Name of the representative(s) of the alleged victims as a character string.

Merits decisions

The following variables are measured at the level of merits decisions. Decisions are treated as distinct if they involve a different article of the American Convention or a different alleged victim. DecisionID can be used to merge with data on Remedial orders. VictimID can be used to merge with data concerning alleged victims. CaseID can be used to merge with other tables in the database.

Table A3: Codebook for Merits table

Variable	Variable description	Coding instructions
DecisionID	The units of analysis in this part of the data	Automatically generated ID variable
CaseID	The case that we can connect the decisions to	ID variable from Case table
VictimID	Who is the victim(s) in this case? Here we connect the victims to the alleged violations.	ID variable from Victims table
AllegedViolationArticleACHR	What article(s) of the American Convention is allegedly violated.	Chosen from drop-down menu of articles of the American Convention
AllegedInRelationTo	Often the alleged violated article(s) are in relation to other articles. If the alleged violated article(s) are not alleged in relation to other articles, we do not code anything here.	Chosen from drop-down menu of articles of the American Convention
AllegedBy	Is it the Commission, Representatives of the applicants, both, or "none of the above" that alleged the violation against the respondent state?	Selection one of the following options: "Representatives of the applicants" "Both" "Not alleged by Commission or representatives"
DecisionOutomce	The IACtHR decision concerning this specific alleged violation for this specific victim	"Violation", "No violation", or "Did not rule on"
Unanimous	Did the Court rule unanimously or not?	Select one of the following options: "Yes" "No"
AcknowledgedByState	Did the respondent State acknowledge responsibility for the alleged violations? We only use the category "refuse" when the State explicitly says that it refuses to acknowledge responsibility for the alleged violation. If there is no information on whether the State acknowledges or refuses responsibility for the alleged violation(s), we code "No info".	Select one of the following options: "Yes", "Partly" "Refuse" "No info"

Remedial orders

The units of analysis are remedial orders. All data concern the judgment on "Reparations and Costs" of the relevant case. RemedyID allows merging this table with data on compliance to track the implementation of each remedial order. DecisionID allows merging with the merits decisions. CaseID can be used to merge with all other tables in the database.

For the MeasureType2 variable, we classify the remedial orders according to the following rules:

- **Legislative changes:** If the State must implement some kind of legislative changes. For example, if the State must make something illegal through law.

- **Prosecution of perpetrator(s)/investigation of crime:** If the State should investigate the crime, or identify, Prosecute, and Punish Those Responsible for the crime.

- **Revoke domestic judgment:** If the Court orders the State to revoke previous domestic judgments.

- **Jurisprudential changes by national courts:** If the Court orders the State to change their jurisprudential practice.

- **Publication/dissemination of the judgment:** If the Court orders the State to publish the judgment or parts of the judgment in for example the newspaper or on the radio.

- **Public acknowledgement of responsibility:** If the Court orders the State to publicly acknowledge responsibility for the violations identified by the case.

- **Practical task (construction, exhume bodies, etc.):** If the Court orders the State to do a practical task. Examples: build a school, provide medical treatments for the victims, build a monument, etc.

- **Education or training programs:** We use this category if the Court orders the State to educate or train state officials. For example, when the Court orders the

State to educate or train the military to respect international doctrines.

- **Judgment as reparation:** When the judgment itself can be considered as reparation.

- **Reinstatement:** If the Court orders the State to reinstate the victim(s)' positions

- **Other forms of executive or administrative action:** If the State has to change administrative or executive practice.

- **Pecuniary damages:** If the Court orders the State to compensate the victim for pecuniary damages.

- **Non-pecuniary damages:** If the Court orders the State to compensate the victim for non-pecuniary damages.

- **Pecuniary and Non-Pecuniary Damages:** If the Court orders the State to compensate the victims for both pecuniary and non-pecuniary damages, and it does not specify which part of the sum should be compensation for pecuniary damages, and which part of the sum should be compensation for non-pecuniary damages.

- **Cost and expenses:** If the Court orders the State to compensate the victims or the representatives of the victims for costs and expenses.

- **Other monetary payments:** We use this alternative if we cannot classify the monetary payments as pecuniary damages, non-pecuniary damages or costs and expenses.

Table A4: Codebook for Remedial Orders table

Variable	Description of variable	Coding instructions
RemedyID	The units of analysis in this part of the data	Automatically generated ID variable
CaseID	The case relevant for the remedial order	ID variable from the case table
DecisionID	The relevant merits decision(s)	ID variable from the merits decisions table
MeasureType1	The title of the remedial order	The title of the remedial order as a string variable
MeasureType2	We categorize the reparations into several categories (see further description in the main text).	Select one of the following categories from a drop-down menu: "Legislative change" "Prosecution of perpetrator(s)/ investigation of crime" "Revoke domestic remedy" "Jurisprudential change by national Courts" "Publication/ dissemination of the judgment" "Public acknowledgment of responsibilities" "Practical task" "Education and training programs" "Judgment as reparation" "Reinstatement" "Other forms of executive or administrative action" "Pecuniary damages" "Non-pecuniary damages" "Pecuniary and non-pecuniary damages" "Cost and expenses" "Other monetary payments"
AmountAwarded	How much money the victim was awarded by the Court in American dollars	The monetary sum in American dollars entered as a number. Leave empty if no money was awarded as part of this remedial order.

xi.

Deadline	Did the Court set a deadline for the implementation of the remedy?	Select one of the following options from a drop-down menu: "Yes" "No"
LengthDeadlineMonths	Time frame allowed before implementation of the remedy ordered by the Court	Time frame in number of months. Leave empty if no deadline was set.
Unanimous	Was the Court's decision on the remedy unanimous? We consult the Court's original documents on the judgments for this information.	Select one of the following options from a drop-down menu: "Yes" "No"

Compliance

The units of analysis are compliance decisions defined as decisions taken with respect to a particular remedial order at a particular compliance hearing. Because compliance decisions are nested within compliance hearings and some variables are constant at the level of the compliance hearings, we include an ID variable for the compliance hearing (ComplianceHearingID) in addition to the ID variable for the compliance decisions (ComplianceDecisionID). RemedyID can be used to merge with the table of remedial orders. CaseID can be used to merge with other variables in the database.

Table A5: Codebook for Compliance table

Variable	Variable description	Coding instructions
ComplianceDecisionID	ID variable for the compliance decisions, which are the units of analysis for this table	Automatically generated ID variable
ComplianceHearingID	ID variable for the compliance hearings that compliance decisions are nested in.	Automatically generated ID variable
CaseID	The case the compliance hearing is part of	ID variable for the case
DateHearing	The date the Court monitors the state's compliance with the Court's rulings	dd.mm.yyyy
JudgesInComplianceHearing	The judges that participated in the Compliance monitoring. This information is not available in the Loyola Law School, Los Angeles case summaries. We therefore look up the Court's original documents on compliance monitoring to gather this data. There are never any ad hoc judges.	Chosen from a drop-down menu of all judges that have been on the Cour
RecusingJudgesInComplianceHearing	A judge might recuse himself/herself from the compliance monitoring. This is usually the case if a judge is a national of the respondent state	Chosen from a drop-down menu of all judges that have been on the Cour relevant" if there are no judges that recused themselves.
AbstainingJudgesInComplianceHearing	A judge might abstain from participating in the compliance monitoring because of "reasons beyond his/her control"	Chosen from a drop-down menu of all judges that have been on the Cour relevant" if there are no abstaining judges in the compliance hear
RemedyID	The remedyID that the compliance decision concerns.	The RemedyID that has been given in the remedy data

ComplianceStatus	Did the Court find that the State fully complied, partly complied or did not comply (pending compliance) with the Court's rulings? If the Court lacks information on the State's compliance with the judgment, we have coded "unclear".	Select one of the following options: "Full compliance" "Partial compliance" "Pending compliance" "Unclear" Code the exact conclusion reached by the court. In particular, "Partial compliance" should only be used when the Court explicitly reaches this conclusion.
Unanimous	Did the Court unanimously decide on the State's compliance status with the Court's judgment? It is often not specified in the Loyola Law School, Los Angeles case summaries or in the Court's original documents whether the judges ruled unanimous or not. If unspecified, we code "no info". However, if a judge had a dissenting, concurring or separate opinion, this is specified in both the case summaries and the Court's original documents.	Select one of the following options: "Yes" "No" "No info"

Dissents

The dissents table contains information about dissenting votes, defined as votes against any type of majority ruling. The units of analysis are the dissenting votes. The judge names can be used to merge with other judge data. DecisionID, RemedyID, ComplianceHearingID, and PreliminaryObjectionID can be used to merge with data concerning the decision the dissenting judge voted against (depending on the type of vote). CaseID can be used to merge with other tables in the database.

Table A6: Codebook for Dissents table

Variable	Variable description	Coding instructions
DissentD	The unit of analysis for the dissent data, i.e. the dissenting votes.	Automatically generated ID variable
CaseID	The case that the judge-vote is part of	ID variable from the case table
JudgeName	Name of the judge who voted against the majority	Name of judge from a dropdown menu of all judges sitting on the case.
TypeVote	In which part of the case did the Judge vote against the majority? "Merits" if the Judge voted against the majority that an article should/should not be considered as violated, or should not be ruled on. "Remedy" if the Judge voted against the majority's decision on the judgment on reparation and costs. "Compliance" if the Judge voted against the majority's decision on the status of the State's compliance with remedies. "Interpretation" if the Judge's votes against the majority's decision to find the interpretation request admissible/inadmissible. "Preliminary objection" if the Judge voted against the majority's decision that they support or object the State's preliminary objection.	Select one of the following options; "Merits" "Remedy" "Compliance" "Interpretation" "Preliminary objections"

xvii

VotingDirection	Did the Judge vote in favor of the respondent State or not?	Select one of the following options: "pro government" "anti government" "Unclear"
DecisionID	The judgment on the merits that the vote concerned.	The DecisionID that has been given under the merits/alleged violation data
MeritVoteDirection	Did the Judge vote that an article should be considered violated, not violated, or should not be ruled upon? We code "not relevant" if the judge's vote did not concern the judgment on merits.	Select one of the following options: "Violation" "No violation" "Not rule on" "Not relevant"
RemedyID	The remedy ordered by the majority of the Court that the vote concerned.	The RemedyID that has been given under the remedy data
RemedyVoteDirection	Did the Judge vote that he/she was against or in favor of a remedy? We code "not relevant" if the judge's vote did not concern the judgment on reparations and costs.	Select one of the following options: "For this remedy" "Against this remedy" "Not relevant"
ComplianceHearingID	The judgment on the State's compliance status that the vote concerned.	The ComplianceHearingID that has been given under the compliance hearing data
ComplianceVoteDirection	Did the Judge vote that the State has fully complied, partly complied or not complied with the Court's judgment? We code "not relevant" if the judge's vote did not concern the judgment on compliance monitoring.	Select one of the following options: "Full compliance" "Partial compliance" "Pending compliance" "Not relevant"
PreliminaryObjectionID	The judgment on the preliminary objection that the vote concerned.	The PreliminaryObjectionID that has been given under the preliminary objections data

		Select one of the following options:
	Did the Judge vote that he/she support or reject the	"Support objection"
PreliminaryObjectionsDirection	State's preliminary objection? We code "not relevant" if the judge's vote did	"Rejects objection"
	not concern the judgment on preliminary objections.	"Unclear"
		"Not relevant"

Opinions

The opinions table contains information about separate, concurring, and dissenting opinions authored by IACtHR. The opinions may concern one or more decisions taken by the Court. The units of analysis are the opinions. DecisionID, RemedyID, ComplianceHearingID, and PreliminaryObjectionID can be used to merge with data concerning the decision that the opinion pertains to.

Table A7: Codebook for Opinions table

Variable	Variable description	Coding instructions
OpinionID	ID variable for the opinions, which are the units of analysis in this table.	Automatically generated ID variable
CaseID	The case the opinion is part of.	ID variable from Case table
DecisionID	Connects the opinion to a specific merits decision	ID variable from Merits table
RemedyID	Connects the opinion to a specific remedy	ID variable from the Remedies table.
ComplianceHearingID	Connects the opinion to a specific compliance hearing	ID variable from the compliance hearing table
PreliminaryObjectionID	Connects the opinion to a specific preliminary objection	ID variable from the Preliminary objections table
JudgeName	Name of the judge with a separate, concurring or dissenting opinion	Name of judge selected from a dropdown menu.
OpinionType	Is the opinion a separate opinion, concurring opinion, or dissenting opinion? In cases where an opinion is partly concurring and partly dissenting, separate entries should be created.	Select one the following options: "Separate" "Concurring" "Dissenting"
OpinionSubject	What is the subject of the opinion?	Select one or more of the following options: "Merits" "Remedy" "Compliance" "Interpretation request" "Procedural matters" "Preliminary objections"

Direction	Is the opinion in favor (pro government) or disfavor (anti-government) of the interests of the respondent State? If the judge's opinion is not concerning the respondent State behavior, but for example what should be the context of international doctrines, we code that the opinion as "neutral".	Select one of the following options: "Pro government" "Anti-government" "Neutral" "Unclear"
MeritsDirection	Does the judge's opinion say that the concerned article is violated, not violated or should not be ruled upon? Even though the Judge voted with the majority in the Court, the Judge could still have a different opinion than the majority of the Court. The Judge could also have an opinion on how an article in the Convention should be interpreted in general, and not say anything about whether the State violated this article or not. In these cases we code "not relevant". We also use the alternative "not relevant" if the judge's opinion did not concern the judgment on merits.	Select one of the following options: "Violation" "No violation" "Not rule on" "Not relevant"
RemedyDirection	Does the Judge say that he/she supports the remedy or rejects the remedy order by the Court? We use the alternative "not relevant" if the Judge's opinion did not concern the judgment on reparation and costs.	Select one of the following options: "For this remedy" "Against this remedy" "Unclear" "Not relevant"
ComplianceDirection	Did the judge find that the State complied fully, partly or did not comply with the Court's ruling? "Not relevant" if the judge's opinion is not concerning the judgment on compliance.	Select one of the following options: "Full compliance" "Partial compliance" "Pending compliance" "Unclear" "Not relevant"

PreliminaryObjectionsDirection	Did the Judge's opinion say that he/she supported the State's preliminary objection, or did the Judge reject the State's preliminary objection? We code "not relevant" if the Judge's opinion did not concern the judgment on preliminary objections	Select one of the following options: "Support objection" "Reject objection" "Unclear" "Not relevant"

Preliminary objections

This table contains information about whether the state filed one or more preliminary objections during the proceedings of the case and, if so, the stated reasons for the objections. The units of analysis are the preliminary objections. CaseID allows merging with other tables in the database.

Table A8: Codebook for Preliminary objections table

Variable	Variable description	Coding instructions
PreliminaryObjectionID	ID variable for the preliminary objections, which are the units of analysis in this table.	Automatically generated ID variable
CaseID	The case that the preliminary objection was filed as part of.	ID variable from the Case table.
PreliminaryObjectionReason	What was the stated reason for the preliminary objection?	The reason for why the state filed a preliminary objection as a character string.

Amici

The units of analysis in this table are *amicus curiae* briefs. CaseID allows merging information about the briefs with other tables in the database.

Table A9: Codebook for Amici table

Variable	Description of the variable	Coding of variable
AmicusBriefID	The unit of analysis for this table	Automatically generated ID variable
CaseID	The case in which this amicus brief was filed to	ID variable from the Case table
Name of Amicus	Name of the Amicus Curiae. This may be an individual person or an organization.	Name of the Amicus Curiae, and other available information as a string variable.

B Reliability

Most cases in our database have been coded by a single coder only (either the first or the third author). This section describes steps taken to ensure consistency in the coding and the reliability tests we have conducted.

B.1 Initial Pilot-Coding and Reliability Assessment

After developing an initial version of the codebook, the two coders (the first and the third author) both pilot coded 20 randomly selected cases for the purpose of establishing a reliability codebook. Reliability was measured by calculating the percentage of cases that both coders had coded identically.

For most variables in the database, there was complete agreement between the two coders while inconsistencies were due to typing errors. For some variables, inconsistencies did, however, reveal ambiguities in the codebook, which were then resolved. In a few instances, we decided to remove variables from the database, because it proved too challenging to establish reliable coding criteria. For instance, we decided not to include information about when the alleged violation occurred because many violations are ongoing over multiple years and it is challenging to establish exact start and end dates.

After this initial pilot coding and reliability exercise, all cases were coded by one of the two coders (from now "the original coders").

B.2 Subsequent Reliability Assessment

After completing the database, we conducted an additional extensive and detailed reliability test. We first trained a research assistant (from now "the reliability coder") previously unaffiliated with the project to recode 40 randomly selected cases for the purpose of assessing reliability. We then calculated the percentage of exact agreement between the two coders. Calculating the exact agreement means for instance that if a case had 44 distinct victims and one of the coders only coded 43 of them, the two coders will be considered to be in disagreement. Finally, we manually investigated the discrepancies between the

reliability coding and the two original coders.

Below, we discuss the results of this reliability exercise variable-by-variable. We organize the discussion by the tables in the database. To summarize, the reliability test resulted in reasonable agreement and most the identified discrepancies are explained by typos or inaccurate data entries by the reliability coder rather than by differences in interpretation.The reliability is reasonably good also for relatively challenging variables. For instance, with regards to our categorization of remedial orders, MeasureType2, it may be difficult to establish whether an order to organize remembrance events should be categorized as a "practical" task or as a "public acknowledgement of responsibility". Nevertheless, even for this type of variable, the inter-coder reliability is 76%. Moreover, because remedial orders can be difficult to classify, we also include a brief description of each order (MeasureType2), allowing users of the database to reconsider our categorization.

For most other variables, the coding is straightforward, and disagreements are only due to data entry errors. For instance, for the outcome of merits decisions, the inter-coder reliability is 95% and when manually investigating the disagreements we only found one data entry error in the original coding (other disagreements being explained by data entry errors by the reliability coder).

The results from the reliability assessment therefore increases our confidence in the reliability of the database. We have corrected all cases where the reliability coding revealed inaccuracies in the original coding of the database.

Cases

A variable-by-variable table of agreement is presented in Table A10. As detailed in the table, the reliability checks revealed few cases of ambiguity or incorrect coding decisions in the original database. Instead, most inconsistencies are explained by the reliability coding being less accurate than our original coding.

Table A10: Reliability check for Cases table

Variable	Agreement	Comments
RespondentState	1	
DatePetition	0.925	In one case, there is an incorrect entry in the reliability coding.
		In one case there are multiple petitions with different dates.
		In a third case, there is a difference between the date the petition was submitted and the date it was received
DateSubmission	0.975	Correct in original coding. One month off in reliability check
DateMeritsRuling	1	
JudgesInMeritsHearing	0.900	Three misentries/typos in reliability coding. One case with one missing judge in original coding
AdhocJudgesInMeritsHearing	0.975	Wrong judge reported in case summary. Corrected by reliability coder
RecusingJudgesInMeritsHearing	0.900	Four mistakes in reliability coding, but all correct in original
AbstainingJudgesInMeritsHearing	0.925	Two mistakes in reliability coding. One mistake in original coding.
DateRemediesRuling	0.925	Three mistakes in reliability coding, but all correct in original
JudgesInRemediesHearing	0.850	Four cases of incorrect data entry in reliability coding
		Two cases of incorrect data entry in original coding
AdhocJudgesInRemediesHearing	0.950	One missing case in reliability coding, correct in original coding
		One case with wrong judge reported in case summary. Corrected by reliability coder
RecusingJudgesInRemediesHearing	0.925	Three mistakes in reliability coding, all correct in original coding
AbstainingJudgesInRemediesHearing	0.900	Three mistakes in reliability coding, all correct in original coding
		One mistake in original coding, corrected in reliability coding
InterpretationRequest	0.925	In two cases, the reliability coder misclassified the actor requesting interpretation
		In one case, the interpretation request was missing from the original coding but corrected by the reliability coder
InterpretationGranted	0.900	In two cases, the reliability coder incorrectly coded that interpretation request was not granted. Correct in original coding.
		In one case, the interpretation was incorrectly coded as not granted in original coding. Corrected in reliability coding
		In one case, the interpretation request was missing from the original coding but corrected by the reliability coder
InterpretationSubject	0.800	In five cases, one of the subjects is incorrectly classified by the reliability coder but correctly coded in the original coding.
		Two cases reveal slight ambiguity between 'assessment of evidence' and the 'merits decision'
		In one case, the interpretation request was missing from the original coding but corrected by the reliability coder
JudgesInInterpretationHearing	0.925	In once case, the reliability coder failed to code one of the judges. Correct in original coding.
		In one case, the interpretation request was missing from the original coding but corrected by the reliability coder
AdhocJudgesInInterpretationHearing	1	
RecusingJudgesInInterpretationHearing	0.925	Three cases where reliability coder failed to code the recusing judge. All correctly coded in original coding
AbstainingJudgesInInterpretationHearing	0.925	Three cases where reliability coder failed to code the abstaining judge. All correctly coded in original coding
AcceptInternationalResponsibility	0.925	Three cases incorrectly classified by reliability coder. All correctly coded in original coding

Victims

Differences in the Victims table are primarily explained by the reliability coder including the victims' next of kin as separate victims even in cases where the next of kin were not identified as such by the judgment if the next of kin received compensation due to the victims being killed. In the original coding, next of kin are only included as victims if their victim status is explicitly recognized by the IACtHR. In cases, where the victim's next of kin are named as separate victims in the judgment, these are registered by name in the original coding of the database but were lumped together as "next of kin" by the reliability coder. In addition, there are two cases with very large numbers of victims in which the reliability coder missed some of the victims named in the judgment.

Users of the database should note that next of kin are only identified in the database when they are named by the Court as separate victims. In cases in which the victim was killed or disappeared and the damages therefore had to be paid out to her family, the next of kin are still not registered as separate victims unless they are explicitly recognized as such by the Court.

Users of the database should also note that the nationality of the victims has only been coded if this information is unambiguous in the case summary. If the nationality is not explicitly stated, we have coded "no info". This is also described in the codebook.

Unclarity concerning the coding of next of kin also reduces the agreement on other variables in this table.

Merits decisions

For the merits decisions, disagreements can be due both to differences in the unique decisions that were identified by the original coders and the reliability coders and to differences in how each decision is coded. To assess agreement affected articles, we compare the overlap between the reliability coder and the original coders by computing the average size of the intersection of the two sets of identified articles in each case divided by the union of the two sets of identified articles (i.e. the Jaccard index). For the remaining variables, we compared the agreement for the articles that were identified both by the original coders

Table A11: Reliability check for Victims table

Variable	Agreement	Comments
VictimName	0.650	Reliability coder has coded the victim's next of kin as separate victims even when they are not named as such by the judgment. Two cases with many victims where the reliability coder has failed to identify all of them
CollectiveVictim	0.625	One data entry error by reliability coder. Other disagreements explained by the incorrect inclusion of 'next of kin'
NaturalPerson	0.650	Members of the YATAMA indigenous group coded as natural person (but also as a collective) in original database, but not in reliability coding. Other disagreements explained by the incorrect inclusion of 'next of kin'
Female	0.800	In two cases, the reliability coder failed to code gender even if gender was identified by pronouns or the use of 'Mr./Mrs.' in the judgment Remaining disagreements are explaining by reliability coder not identifying all victims.
CitizenOfRespondentState	0.075	In cases where nationality is not mentioned in the judgment, the original coding is 'No info'. This is described in the codebook. The reliability coder has nevertheless coded these as nationals of the respondent state.
RepresentedBy	0.825	Six cases in which the reliability coder entered incorrect information, but where the original coding is correct. One case where reliability coding is correct, but representative is missing in original coding.

and the reliability coding. The results from the reliability checks are summarized in Table A12.

The reliability check has not revealed discrepancies that are due to ambiguities concerning how some variables are to be coded, which strengthens our confidence in the original coding. Instead, most disagreements concern cases where the reliability coder has simply been less accurate than the original coders.

Remedial orders

The reliability assessment for the Remedial orders table, similarly to the Merits table, needs to consider both the overlap in the orders that were identified and how each remedial order is coded.

To assess the overlap in identified remedial orders, we created two sets of combinations of the CaseID and the MeasureType2 variable used to classify orders and calculate the Jaccard index as the size of the intersect of these two sets divided by the size of the union of the two sets. The Jaccard index may be interpreted as the percentage agreement between the two sets and is .65. This relatively low level of disagreement is explained primarily by the reliability coder identifying fewer remedial orders than the original coding. In the

Table A12: Reliability checks for Merits table

Variable	Agreement	Comments
IACHRarticle	0.846	There are discrepancies for 20 of the recoded cases. Discrepancies in 16 of these cases are due to the reliability coder not adding all the decisions in the case, while there are four cases with missing decisions in the original coding.
VictimID	0.790	Some discrepancies are due to the reliability coder grouping some victims together, for instance as "next of kin" while the original coders coded the full list of victims. In other instances, there are multiple decisions pertaining to the same article, but involving different victims, where the reliability coder has not coded all decisions. We recoded 20 of the decisions with disagreements and only found problems with the original coding for three decisions.
AffectedInRelationTo	0.853	Almost all discrepancies are due to the reliability coder coding too few articles. We recoded 20 of the decisions with disagreement and only found problems with the original coding for three decisions.
AllegedBy	0.749	We recoded 10 cases with disagreements and only found problems with the original coding in one case. It is unclear why inaccuracies in the reliability coding have occurred as the actors alleging each violation are clearly listed in the case summaries.
DecisionOutcome	0.953	We recoded all 15 cases with disagreement. In a single case the original coders had coded "no violation" instead of "not rule on". In all the remaining cases, the reliability coder failed to accurately code the Court's decision.
AcknowledgedByState	0.787	All except three instances of disagreement are due to the reliability coder coding "No info" even if there is information that the state accepted or refused responsibility. The remaining three cases are all correctly coded by the original coders but incorrectly coded by the reliability coder.
Unanimous	0.947	We recoded all cases with disagreement (see also results for the Dissents table). The original coding proved to be correct for all cases with disagreements.

40 re-coded cases, the reliability coder identified a total of 424 remedial orders compared to a total of 473 remedial orders in the original coding.

To investigate the source of these discrepancies, we aggregated both the original coding and the reliability coding to the level of CaseID x MeasureType2 and sampled 20 cases in which the count differed between the two datasets. In 12 of these 20 cases, the disagreements where explained by the original coding being more disaggregated. For instance, the original coding disaggregated monetary awards where there are multiple victims to allow tracking how much money each victim is awarded. Similarly, for orders to investigate or prosecute crimes, the original coding distinguishes between different crimes/perepetrators etc. to allow capturing that the state may comply with some, but not all of the orders in a category. By contrast, the reliability coder has been more willing to aggregate these instances into single orders.

In the remaining 8 cases of disagreement concerning the number of a measure type within a case, the disagreements can be explained by ambiguity concerning how an order is best classified. For instance, one order to identify and exhume bodies has been coded as "Practical task (construction, exhume bodies, etc.)", but as orders to conduct a "Prosecution of perpetrator(s)/investigation of crime" by the reliability coder. While the original coding classifies an order to organize a remembrance event as a practical task, the reliability coder coded the same order as a public acknowledgment of responsibility. Precisely because not all remedial orders are easily classified, we also supply the MeasureType1 variable, which contains a qualitative description of each order.

For the remaining variables in this table, we aggregated the data to the level of CaseID x MeasureType2 and calculated agreement on the mean level on each variable. The results are reported in Table A13. The table shows that while the agreement on the number of each type of remedial order in each is only .76, there is a high level of agreement on the other variables. The main source of disagreement between the reliability coder and the original coding is thus how many remedial orders are identified for each case. Similarly as the other tables, we found that for most cases of disagreement, the original coding was accurate. This strengthens our confidence in the original coding.

Table A13: Reliability checks for Remedial Orders table, after aggregating to level of CaseID x MeasureType2

Variable	Agreement	Comments
Count of each MeasureType2 in case	0.761	Disagreements primarily concern cases with monetary awards or orders to prosecute perpetrators and are in part explained by the original coding disaggregating cases where different victims are offered different awards or where there are orders to investigate/prosecute perpetrators from multiple events. By contrast, the reliability coder has tended to aggregate these orders. In addition, some remedial orders have been classified differently by the reliability coder and the original coding. For instance the reliability coder has coded an order to "organize a remembrance" as "public acknowledgment", while the same order has been coded as a practical task in the original coding. Users of the database should be advised that distinguishing between different types of orders is not always clear cut and for this reason we also include a qualitative description of the order (see also discussion in main text).
AmountAwarded	0.823	Disagreements mostly concern complex cases with multiple victims. In total 37 disagreements were identified. In 5 of these, the judgments are ambiguous and it is challenging to establish the exact amount awarded. In 25 cases, the original coding was correct and mistakes had been made in the reliability coding. In 7 instances, the reliability coder was correct, while the original coding was inaccurate.
Deadline	0.866	27 orders with disagreement, of which 21 were correctly coded in the original coding, but incorrectly coded in reliability coding. 6 orders are correctly coded in reliability coding, but incorrectly coded in the original coding.
LengthDeadlineMonths	0.842	19 disagreements are due to disagreement about whether a deadline was set (see variable above). Of the remaining 12 cases of disagreement, 9 orders were correctly coded in the original, while the reliability coding was correct in 3 instances.
Unanimous	0.919	17 orders with disagreement, of which 13 are correctly coded in the original coding and incorrect in the reliability coding. 4 orders are coded correctly in reliability coding and incorrect in original coding.

Compliance

Results from the reliability checks for the Compliance table are reported in Table A14.

For the variables "Compliance status" and "Unanimous", we calculated agreement at the level of the CaseID-MeasureType2-DateHearing level for those remedial orders identified by both the reliability coder and the original coders (see above for a discussion of discrepancies in the identification of remedial orders). While there is a high level of agreement concerning whether the decision on compliance status was unanimous, there are some discrepancies in the coding of the compliance status. These discrepancies appear to primarily be cases where the Court discusses challenges in determining the compliance status, but concludes that the state has fully complied or still not complied and where the reliability coder has coded "partial compliance". Consistent with the codedbook, the original coders have followed the Court in coding these either as "full compliance" or "pending compliance" and have reserved the "partial compliance" category for instances where this is the conclusion reached by the Court. Users of the database should take note of this coding rule.

For the DateHearing variable, we compared hearing dates identified for each case. There are some discrepencies, which are either due to cases having a larger number of hearings and where information on a single hearing is missing or the reliability coder registering incorrect dates for some of the hearings (e.g. June instead of July). The first source of discrepancy is more problematic as it occurs also in the original coding.

For coding of judges at each compliance hearing, we compare the coding at the level of CaseIDxDateHearing. For the judge variables, the agreement is generally high and discrapencies appear to be due to data entry errors.

Dissents

For the dissents table, we first compared the dissenting votes identified by the original coders and the reliability coder. The reliability coding did not identify any dissenting votes that were not identified in the original coding. However, the original coding identified dissenting votes by three judges in two different cases that were not included in the

Table A14: Reliability checks for Compliance table

Variable	Agreement	Comments
ComplianceStatus	0.776	Disagreements primarily concern cases as "Partial compliance" by the reliability coder even if the relevant orders are recognized as fully complied with or deemed to have not yet been complied with by the original coders. As explained in the codebook, the original coders only coded outcomes as "partial compliance" for cases where this conclusion was explicitly reached by the Court. An outcome may thus be coded as "full compliance" even if not everyone would agree that the outcome is perfect as long as the Court considers that the state has fully complied. This coding rule has been consistently applied by the original coders, but not by the reliability coder.
Unanimous	0.959	There are differences in the coding of eight remedial orders from two different hearings. In one of these, the reliability coder has incorrectly coded "no" for unanimous decisions. In the other hearing, there were no values entered by the original coder.
DateHearing	0.655	There are discrepancies in the registered dates in 10 cases. Of these four are cases in which the reliability coder entered the wrong date for one of the hearings, two are cases in which a single hearing was missing from the reliability coding, and four are cases in which a single hearing was missing from the original coding.
JudgesInComplianceHearing	0.906	Three hearings with one missing judge in original coding and two hearings with one missing or incorrectly coded judge in the reliability coding.
RecusingJudgesInComplianceHearing	0.954	Two cases where the recusing judge was missing from the reliability coding, but correctly coded in original coding. One case where the original coding had incorrectly coded the recusing judge as abstaining.
AbstainingJudgesInComplianceHearing	0.908	Three cases where the abstaining judge was missing from reliability coding and Three cases where the abstaining judge was missing in the original coding (in one of these, the abstaining judge was coded as recusing).

reliability coding. Thus, although the reliability coder failed to register dissenting votes in three cases, the reliability check does not suggest that missing dissents is a problem for the original coding of the database.

For the dissenting votes coded by both the reliability coder and the original coding, we compared the coding of the type, content, and direction of the votes. The results are reported in Table A15. With one exception, the disagreements between the two coders are due to the reliability coder missing some of the decisions the dissenting judges voted against. Typically, the dissenting judge has voted against both merits and remedial decisions, and the reliability coder has only coded a subset of all the decisions. In all these cases, the original coding is, however, correct, suggesting the original coding has been able to register all dissenting votes.

In one case, the original coding has coded a dissenting vote as being in favor of a violation even if dissented against the violation finding. This vote was coded correctly by the reliability coder. This error account for the lack of full agreement on the "DissentContent" and "DissentDirection" variables.

Table A15: Reliability checks for Dissents table

Variable	Agreement	Comments
TypeVote	0.556	Four cases where the reliability coding had failed to code all decisions that the dissenting judge dissented against. Original coding is correct in all cases.
DissentContent	0.889	One case where the original coding was incorrect.
DissentDirection	0.889	One case where the original coding was incorrect.

Opinions

For the opinions table, we first counted the distinct number of judge-opinions identified in each case by the original coding and the reliability coding. The agreement was 0.975. The only discrepancy was for one case, in which the reliability coder had failed to register one of the dissenting opinions.

Next, we compared the coding of each opinion identified by both the original coding and the reliability coder. The results are displayed in Table A16

Table A16: Reliability checks for Opinions table

Variable	Agreement	Comments
OpinionType	0.926	Four incorrect entries in reliability coding, which were all correct in the original coding.
OpinionSubject	0.574	There were 23 disagreements, 3 disagreements concern complex opinions where both coders may be considered at least partially correct 16 disagreements concern cases where the original coding was correct and the reliability coding was inaccurate. 4 disagreements concern cases where the reliability coder was correct and the original coding was inaccurate. Most disagreements either concern opinions with multiple subjects, in which the reliability coder failed to enter all subjects, or relatively challenging opinions where the reliability coder coded "Unclear", but the subject was established in the original coding.
Direction	0.740	There are 14 discrepancies, which in part are related to difficulties in establishing whether concurring opinions are against the interest of the respondent state or should be coded as "neutral". When reassessing the cases with discrepancies, we found that the original coding was correct in 11 of these cases, while the reliability coding was correct in 3 cases. Users of the database should, however, note that the direction can be challenging to code for concurring/separate opinions.
MeritsDirection	0.731	7 cases of disagreement concerning concurring opinions that relate more to the reasoning of the Court than the disposition of the case. The original coding has only coded a specific direction in cases where the opinion discusses the disposition of the case, rather than just the doctrine. The reliability coder did not apply this rule consistently.
RemedyDirection	0.273	8 cases of disagreement of which the 7 are cases where the reliability coder has coded "For this remedy", while the original coder has coded "Not relevant". Similarly to for merits decisions, the original coding only codes the direction when the opinion argues for or against a specific remedial order. The reliability coder did not apply this rule consistently. In one case, the reliability coder coded an opinion as being in favor of a remedy, although it argued against the remedial order. This opinion was correctly coded in the original coding.
ComplianceDirection	1	

Preliminary objections

For the Preliminary Objections table, we compared the number of distinct objections identified by the original coding and the reliability coder for each case. The results are reported in Table A17 and shows an 87.5 % agreement. We looked up the cases with disagreements and found that in all instances the original coding was correct, while the reliability coder had failed to enter one or more of the objections filed in the case.

Table A17: Reliability checks for Preliminary Objections table

Variable	Agreement	Comments
Preliminary objections	0.875	In all cases of disagreement our original coding was correct, while there were mistakes in the reliability coding.

Amici

For the *amici* table, we compared the number of *amicus curiae* briefs identified by the original coding and the reliability coder for each case. The results are reported in Table A18 and shows a 90% agreement. Three of the cases of disagreement are cases where multiple briefs had been submitted and one of the coders failed to register one of them. In one case, the original database had not coded the only brief submitted in the case.

Table A18: Reliability check for Amici table

Variable	Agreement	Comments
Amicus curiae submissions	0.900	In two cases, the reliability coder missed one of the briefs submitted. In two cases, briefs were missing from the original coding.

C   Comparison with Hillebrecht's Compliance with Human Rights Tribunals (CHRT) Dataset

Most of the data in our database has not previously been made publicly available in a systematic format. However, the "Compliance with Human Rights Tribunals (CHRT) Database" made available by Hillebrecht (2014a,b) contains information about compliance with remedial orders from 65 IACtHR judgments. To assess construct validity, this section compares the categorization of different types of remedial orders and the coding of compliance status for the 65 available cases.

The only identifying information in the CHRT database are the names of the IACtHR cases and the variable "mandatetype" which distinguishes between "financial reparations", "symbolic measures", "retrials and accountability", "measures of non-repetition", and "individual measures". We are therefore not able to merge the two datasets at the level of the individual order. However, by reconstructing Hillebrecht's categorization of remedial orders, we can compare the distributions of compliance orders falling within each category and compliance rates for the different types of measures. This exercise also illustrates how users of the database may use our fine-grained categorization of remedies to construct their own categorizations.

We first mapped values on our MeasureType2 variable to Hillebrecht's categorization using the instructions in her codebook[1] and in Hillebrecht (2014a, 50–51). For the values on MeasureType2 relevant to cases in both databases, we use the mapping in Table A19 to reconstruct Hillebrecht's mandatetype variable based on our MeasureType2.

The distributions of different types of orders in CHRT database and in our database can be assessed by comparing the upper panel in Figure A1 – displaying the distribution in the CHRT database – to the two lower panels in the same figure – displaying the distribution in our database. As can be seen, the distributions are very similar except for the category "Financial reparations" for which our database has a much larger number of distinct orders. The reason for this discrepancy is that we code separately orders that

---

[1]https://courtneyhillebrecht.files.wordpress.com/2016/02/chrt_codebook.pdf (retrieved March 8th, 2020).

Table A19: Mapping our MeasureType2 variable to Hillebrecht's mandatetype variable

mandatetype from Hillebrecht	Our MeasureType2
1 Financial reparations	Costs and Expenses, Pecuniary and Non-Pecuniary Damages, Non-Pecuniary Damages, Other monetary payments, Pecuniary Damages
2 Symbolic measures	Publication and dissemination of the judgment, Public acknowledgement of responsibility, Judgment as reparation
3 Retrials and accountability	Prosecution of perpetrator/investigation of crime, Revoke domestic judgment, Jurisprudential changes by national courts
4 Measures of non-repetition	Legislation, Other executive or administrative task, Education and training
5 Individual measures	Practical task (construction, exhume bodies, etc), Reinstatement

concerns a different victim or that we code different types of compensation (e.g. pecuniary and non-pecuniary damages) separately. For compliance scholars, this additional level of detail is useful because a state may compensate some but not all the victims in a case. In addition, the more disaggregated coding will allow researchers to explore a range of other research questions, such as what determines the size of IACtHR monetary awards.

The CHRT database contains a dummy indicator for whether each remedial order had been complied with to the "Inter-American Court's satisfaction. If states have partially complied with a particular obligation, the value is 0. A score of 1 indicates complete compliance with that obligation."[2] The codebook further notes that the data is right-censored and that the recorded compliance status reflects the status when the case was coded and that the coding occurred during the 2008-2010 period. Because there is no exact censoring date, we cannot establish exactly for which point in time the compliance status are accurate. However, because our database keeps track of the compliance status of each order at each compliance hearing, it is possible to compare the two at different

---

[2]https://courtneyhillebrecht.files.wordpress.com/2016/02/chrt_codebook.pdf(retrieved March 8th, 2020)

points in time. In Figure A1, we offer two comparisons which we consider useful. In the mid-panel, we show the highest compliance status achieved by the end of 2010 according to our database. In the lowest panel of the same figure, we show the highest compliance status achieved to date, according to our data.

Comparing the share of orders in each category that are coded as complied with by Hillebrecht and were "fully complied with" by 2010 according to our data shows a high level of agreement. Again, the main discrepancy concerns financial reparations for which our database codes a higher share of orders that have been fully complied with. A likely explanation is that in cases where states comply with some remedial orders but not others, the CHRT data will record that the order for "financial reparations" has not been complied with, whereas our database will record some cases of full compliance and some cases of partial, pending, or unclear compliance.

Figure A1 also illustrates two further points. First, because our database separates between four different compliance outcomes, we provide more information about the status of orders that receive the value of 1 on the compliance dummy in the CHRT data. Users of our database can thus distinguish cases where the Court has determined that the state has not yet complied from cases where it says that the compliance status is still unclear.

Second, Figure A1 shows the importance of tracking compliance outcomes over time. Comparing the mid-panel to the lowest panel shows that since 2010, the share of orders with "partial compliance" and "unclear" compliance has slightly decreased while the share of cases with "full compliance" and "pending compliance" have slightly increased. As time passes, states are able to fully comply with more measures, and where compliance is not achieved the IACtHR becomes more willing to call them out in compliance hearings by declaring that compliance with an order is still pending.

To further compare the compliance coding in our database with what is currently available in the CHRT data we calculated shares of cases that had received the status of full compliance by 2010 by case and "mandate type" in the two databases. Figure A2 plots the compliance rates from both databases against each other (using a 0.1 jitter). The correlation is .65, which – considering uncertainty concerning the exact censoring

date in the CHRT data and the different ways orders are aggregated in the two datasets

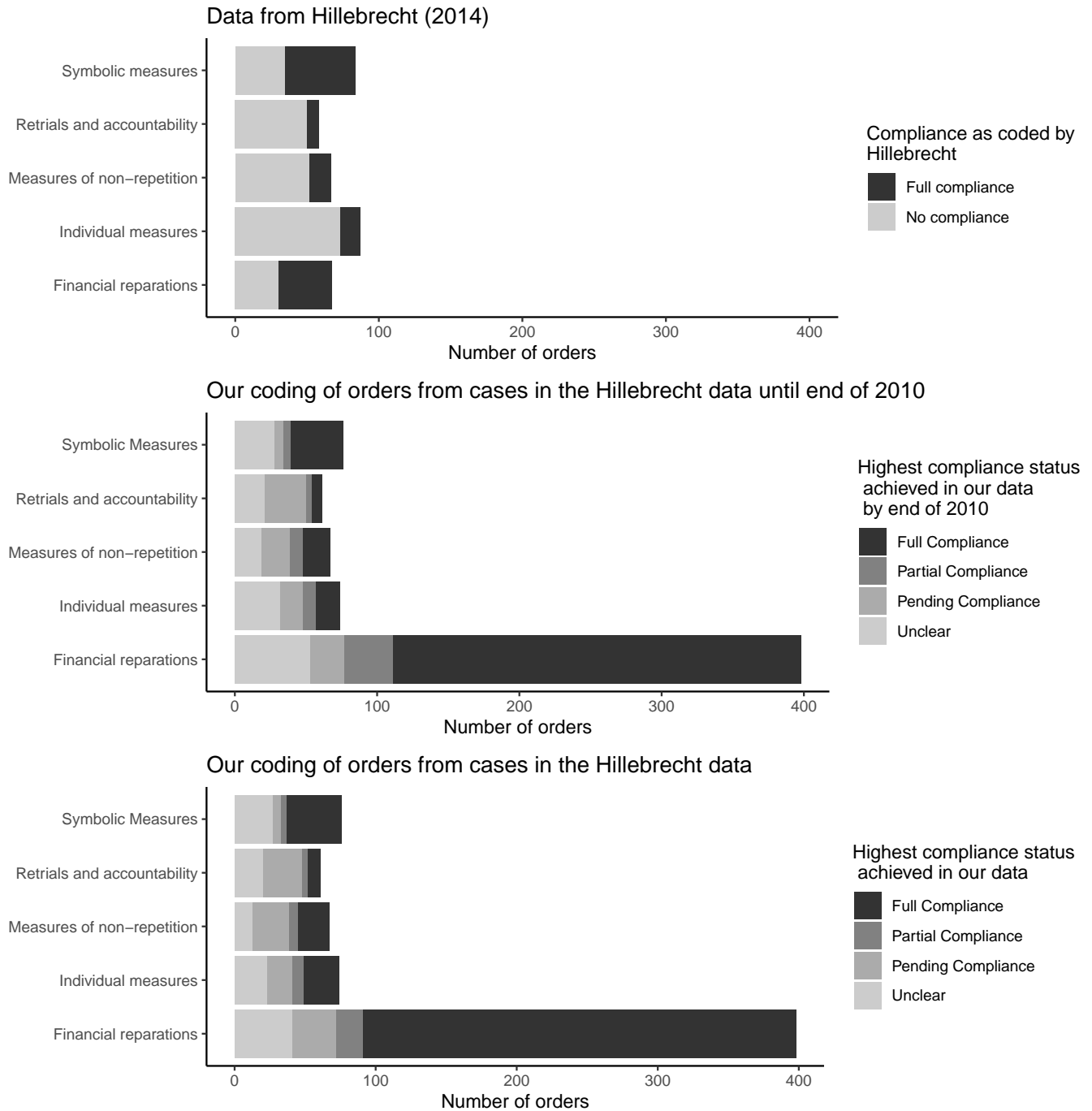– must be considered quite high.

Figure A1: Remedial orders included in Courtney Hillebrecht's CHRT database. Comparison between Hillebrecht's and our coding.
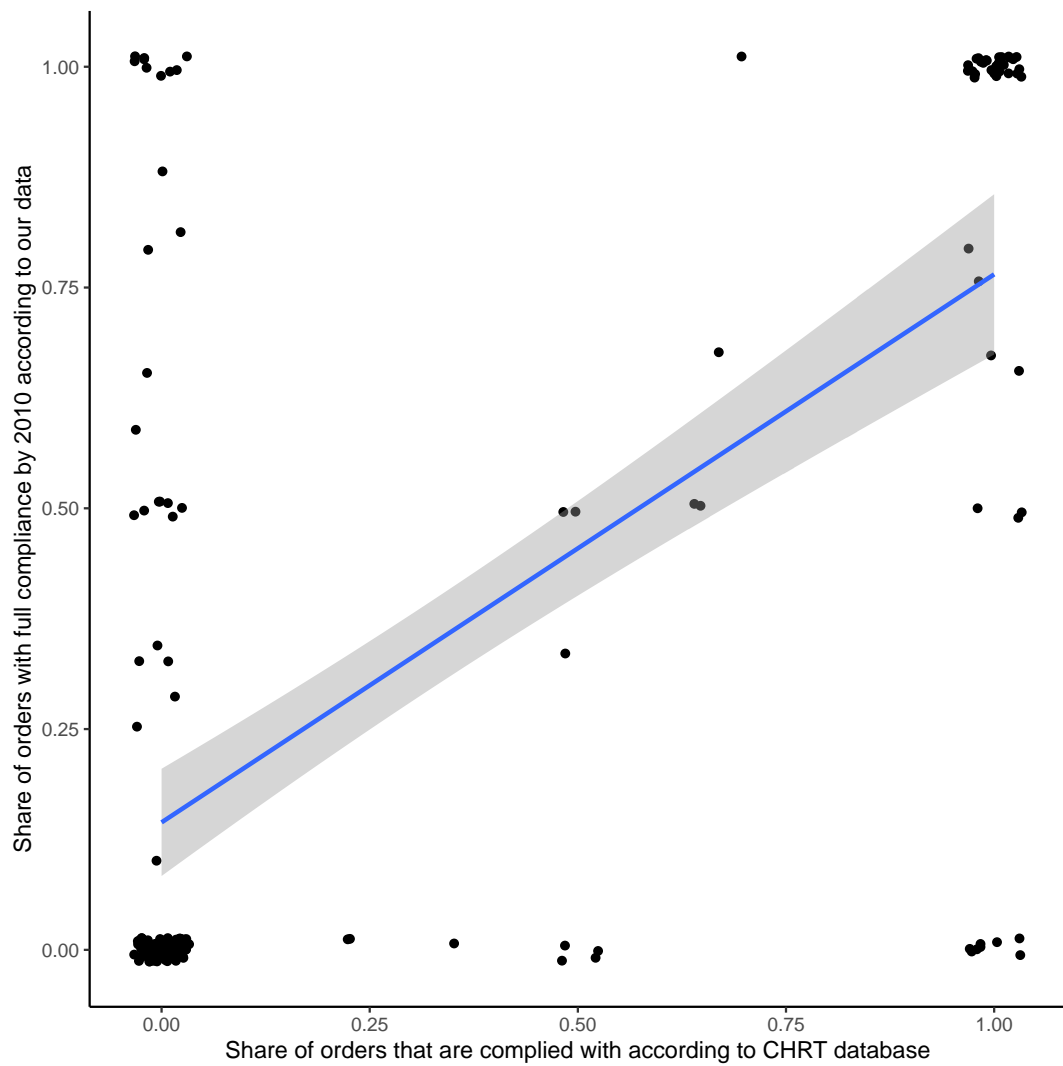
Figure A2: Compliance rate by "mandate type" and case in our data vs. the CHRT database. Positioning with 0.1 jitter.

References

Hillebrecht, Courtney. 2014*a*. *Domestic politics and international human rights tribunals: the problem of compliance.* Cambridge: Cambridge University Press.

Hillebrecht, Courtney. 2014*b*. "The power of human rights tribunals: Compliance with the European Court of Human Rights and domestic policy change." *European Journal of International Relations* 20(4):1100–1123.