The Conditional Effectiveness of Soft Law: Compliance with the Decisions of the Committee against Torture

Andreas von Staden

Abstract

The article examines the record of compliance with the UN Committee against Torture’s decisions in individual complaints cases. Theoretically, I expect that compliance will be the outcome of a combination of normative and rationalist factors: States committed to human rights protection will comply even in the absence of enforcement but only as long as compliance costs remain relatively low. Using a data set covering all adverse decisions issued until 2018 and information on their compliance status, I employ fuzzy-set qualitative comparative analysis to identify necessary and sufficient conditions for the outcomes of compliance and noncompliance. The analysis reveals that the conditions tested—liberal democracy, nonuse of political terror, violation type, and strong civil society—are in part individually (near-)necessary and jointly sufficient for compliance, while the presence of their complements is consistent with noncompliance. The Committee against Torture thus appears to be able to elicit (some) state compliance with its decisions, if mostly only with respect to certain kinds of states and select types of violations.

Keywords UN Committee against Torture · Individual communications procedure · Compliance · Legal and nonlegal bindingness · Liberal democracy · Qualitative comparative analysis (QCA)

Introduction

The United Nations Committee against Torture (CAT) is the second-most-active UN human rights treaty body—after the Human Rights Committee—in dealing with individual complaints. Since 1989, CAT has registered 1,068 complaints and has issued decisions on the merits on close to 400 of them, with a backlog of 219 complaints pending consideration as of April 2021 (Committee
against Torture 2021: para. 47). However, while the committee and the 1984 Convention against Torture have figured prominently in the political science literature on international human rights protection, the effectiveness of CAT’s individual communications procedure (ICP) in eliciting state compliance with its decisions has not been examined. Extant research has investigated why states commit to the convention and/or committee’s ICP (Goodliffe and Hawkins 2006; Hathaway 2007; Hafner-Burton et al. 2008; Vreeland 2008; Simmons 2009; Hollyer and Rosendorff 2011; Sandholtz 2017; Hong and Uzonyi 2018), whether such commitment improves aggregate human rights protection on the ground (Hathaway 2002; Hafner-Burton and Tsutsui 2005; Neumayer 2005; Powell and Staton 2009; Hill 2010; Cole 2012; Conrad and Ritter 2013), the quality of state participation in state reporting procedures (Creamer and Simmons 2015), and the latter’s impact on the human rights situation in target countries (Creamer and Simmons 2019; McQuigg 2011). Commitment to CAT’s ICP acts as either a dependent or independent variable that is taken to indicate deeper or more sincere commitment than treaty ratification alone, but its output’s effectiveness has not been examined more closely.

This article offers a first exploratory foray into the analysis of compliance with the committee’s decisions in individual complaint cases. Theoretically, I expect that compliance with CAT decisions results from an interaction between normative and rationalist considerations on the part of respondent states. Although legally nonbinding, I theorize that CAT decisions generate a form of nonlegal bindingness for states that appear to be committed generally to human rights protection. However, this normative effect’s efficacy in eliciting remedial changes is conditional on the expected costs of compliance, that is, if compliance becomes too costly, and expected costs of noncompliance remain negligible, states increasingly opt for partial, deficient, or outright noncompliance. The normative element explains why states tend to comply with CAT decisions even in the absence of meaningful enforcement, whereas the rationalist element points to the scope condition of acceptable compliance costs.

Because I expect qualitative differences between respondent states and between types of violations/decisions to matter for compliance and noncompliance, I use fuzzy-set qualitative comparative analysis (fsQCA) to test my expectations. Employing a comprehensive data set of all adverse decisions rendered by CAT until the end of 2018 and information on their state of compliance, I show that compliance with CAT decisions is driven by certain value orientations of respondent states as well as implementation costs: When the latter are relatively low, as they are in the frequently found conditional violations of the non-refoulement norm, liberal democracies in particular comply with CAT decisions despite the absence of legal bindingness and meaningful enforcement. When other types of violations and non-democracies are involved, however, noncompliance becomes the dominant outcome. The results suggest that CAT decisions can generate some compliance pull for some states but fail to do so when compliance costs increase and/or sincere normative commitment on the part of respondent states is lacking.
The Conditional Effectiveness of Soft Law: Compliance with…

**CAT’s Individual Communications Procedure**

The 1984 Convention against Torture and Other Cruel, Inhuman, or Degrading Treatment or Punishment entered into force on June 26, 1987, after 20 states had ratified it. The convention includes inter alia a definition of torture (Article 1 [1]), stipulates a *non-refoulement* norm with respect to threats of torture (Article 3), obligates state parties to criminalize all acts of torture (Article 4), and outlaws “other acts of cruel, inhuman, or degrading treatment or punishment which do not amount to torture” (Article 16 [1]). Article 17 (1) establishes the Committee against Torture as a monitoring body composed of 10 independent experts. As of December 22, 2021, the convention counts 173 state parties (United Nations Treaty Collection 2021).

Article 22 establishes CAT’s individual communications procedure. As with the other treaty bodies, this procedure is optional and requires a separate declaration of acceptance. It has been active since the convention entered into force (United Nations Treaty Collection 2021: Note 1), with 69 states accepting it so far (≈ 40.4% of all state parties) (Committee against Torture 2021: para. 42). While Article 22 speaks of “communications” and the committee’s “views,” since the adoption of a set of revised rules of procedure in 2002—covering the bulk of its ICP output—the committee refers to them as “complaints” and “decisions,” respectively (Committee against Torture 2002: Annex X, Rules 96 *et seq.*). The convention does not define the decisions’ legal status; however, as in the case of the other treaty bodies, the dominant position in state practice and the literature is that they are not legally binding (van Alebeek and Nollkaemper 2012: 373; Monina 2019: 631) and are rather a type of “soft law” (Cerone 2016: 23; Shelton 1997: 125).

Also in 2002, the committee created the role of *rapporteur* for follow-up, with a “mandate, *inter alia*, to monitor compliance with the committee’s decisions” (Committee against Torture 2002: 220). When states adopt appropriate remedial measures, CAT will end its follow-up procedure with a note of “satisfactory” or “partially satisfactory” resolution; otherwise, follow-up continues. The measures expected for a finding of satisfactory resolution are typically those indicated in the committee’s decision (Monina 2019: 633–636) and can include an applicant’s non-removal (or safe return, if already deported) and the granting of a residency permit; the investigation, prosecution, and punishment of those responsible for a violation; compensation for “material and psychological harm suffered”; release from detention; and legislative changes.\(^2\)

---

1. Whether the Spanish Supreme Court’s 2018 decision that CEDAW decisions are binding on Spain represents a “Copernican turn” (Kanetake 2019: 590) that will be followed by governmental actors in other states remains to be seen.

2. See, e.g., *Dar v. Norway* (CAT/C/38/D/249/2004), para. 16.4; *Sonko v. Spain* (CAT/C/47/D/368/2008), para 10.9, *E.N. v. Burundi* (CAT/C/56/D/578/2013), para. 9, *Ramírez Martínez et al. v. Mexico* (CAT/C/55/D/500/2012), para. 19.
The lack of legal bindingness has been repeatedly singled out as an apparent reason for the modest rates of compliance with treaty body views (Borlini and Crema 2019; Ulfstein 2014: 253; Heyns and Viljoen 2002: 29, 33), but legal status—understood as the presence or absence of legal bindingness—is neither necessary nor sufficient for either compliance or noncompliance. After all, a non-trivial number of committee decisions have been complied with despite the lack of legal bindingness, while many regional human rights courts’ legally binding judgments remain in a state of partial or full noncompliance (Hillebrecht 2014: chap. 3; Hawkins and Jacoby 2010). Legal status is just one factor among several that may affect compliance in its interaction with other causal influences. More importantly, bindingness and obligation are not exclusive to legal norms and can induce compliance outside the law. As Kal Raustiala (2005: 611) notes, “factors that push states to comply with [legally binding treaties] often apply, albeit more weakly, to [soft law] as well.”

I expect that the interaction of two factors will determine compliance and noncompliance with CAT decisions: The first is the normative factor of the extent to which respondent states accept such decisions as restraints on their freedom to act, that is, as binding them. The second, rationalist factor foregrounds the expected political and material costs of compliance (and noncompliance) which are expected to constrain and set limits on the efficacy of the normative factor in producing compliance.

**Normative Factor: (Nonlegal) Bindingness**

Discussions of the legal status of treaty body views often misleadingly infer that the absence of legal bindingness must mean that they are nonbinding. Strictly speaking, this is a non-sequitur originating from a dichotomous conceptualization that falsely views bindingness as an exclusively legal quality (see, e.g., Broude and Shereshevsky 2021). However, being “law” or “legal” is not a constitutive element of bindingness. Instead, the ordinary meaning of “to bind” is simply “to put under an obligation” or “to exert a restraining or compelling effect,” with “to constrain with legal authority” being only one specific meaning, not the only one (Merriam-Webster’s Collegiate Dictionary 10th ed. 1993: 114). Thus, saying that a norm, decision, or commitment is binding means that what is being decided or agreed upon is no longer subject to the unrestricted discretion of the parties involved. In other words, being bound by a norm, decision, or commitment implies that discretion is constrained by it. When viewed this way, a legally nonbinding norm or decision can be nonlegally binding politically, socially, or morally and generate concomitant expectations of compliance. If legal bindingness means that “a legal reasoner should … take [that norm’s] content into account whenever it is relevant, giving it the weight it deserves” (Sartor 2008: 217), then nonlegal bindingness entails the same in the realm of political, social, or moral obligations.
This hardly comes as news to constructivists who have long disassociated norms’ behavioral significance from their legal status. Positing that norms, regardless of legal status, “prescribe (or regulate) behavior” by expressing “collective expectations for the proper behavior of actors with a given identity” (Katzenstein 1996: 5) is another way of saying that such actors are, at least to some extent, constrained, or bound, by such norms: “Claims of ‘oughtness’ [that are a constitutive element of a norm] are not understood as optional, such that a social community would react in the same way whether individuals complied or did not comply with a given norm. Should one opt not to comply with a norm, we should expect a reaction from the social group to signal disapproval with their deviant behavior” (Jurkovich 2020: 695). Without assuming some such binding effect, and if norms could be ignored without any adverse consequences whatsoever, the stipulated behavioral link between identities and norms would lose much of its explanatory leverage.

Likewise, the notion of soft law takes its cue from the realization that soft law commitments “are not simply politics” and generate at least “hortatory obligations” (Guzman and Meyer 2010: 172) that operate and may elicit effects in ways similar to law (Cerone 2016: 16–17; Shelton 2000). Similarly, in the legalization literature, the absence of intent to be bound legally results in “low obligation,” but not its absence altogether (Abbott et al. 2000: 410), and even “soft commitments may … implicate the legal principle of good faith compliance” (Abbott et al. 2000: 412).

As Oscar Schachter (1977: 300) noted long ago, the fact that “noncompliance by a party [with a legally nonbinding agreement] would not be a ground for a claim for reparation or for judicial remedies … is quite different from stating that the agreement need not be observed or that the parties are free to act as if there were no such agreement.” Many governance arrangements are characterized in terms of nonlegal bindingness. For example, the Organization for Security and Cooperation in Europe (OSCE) has been described as rendering “politically, but not legally, binding decisions” (Greer et al. 2018: 4), and political bindingness has been asserted with respect to commitments at G7 and G20 summits as well as Arctic Council output (Kirton 2019: 6; Rottem 2020: 56). Raustiala, epitomizing the difference between law and soft law as one between contracts and pledges, put it thus: “Contracts create legally binding obligations for states, while pledges create only political or moral obligations” (2005: 586).

Political and/or moral obligations need not be any less consequential than legal ones, though. The real issue of theoretical and empirical importance is what conditions make normative pronouncements “more likely to generate a sense of obligation, and corresponding behavior change” (Finnemore and Toope 2001:749). Adhering to legitimacy criteria, including conformity with other rules and values,
is often identified as a source of obligation (ibid.: 749–750) that applies to law and soft law alike (Höflinger 2020: 669–671; Karlsson-Vinkhuyzen and Vihma 2009; Franck 1990). Brunnée and Toope, in examining legal obligations subject to Fuller’s “conditions of legality,” acknowledge that the bindingness of soft law sometimes may exceed that of formal law: “‘Soft’ norms may sometimes possess more obligatory force than norms derived from formal sources of law. […] When norms are rooted in shared understandings and adhere to the conditions of legality, they generate fidelity” (Brunnée and Toope 2010: 51).

The constraining effect of legally nonbinding decisions can therefore be understood as a function of shared understandings and internalized commitments that are linked to state identity. States that are committed to human rights protection and that view CAT’s ICP decisions as the legitimate outcome of a freely accepted procedure that deserves deference are likely more inclined toward complying with these decisions than states that lack such commitment. Thus, bindingness is a variable: Not all states feel similarly constrained by treaty body output, nor do the same states necessarily feel similarly bound across different treaty bodies, nor across different types of decisions by the same body, as they may view approaches to and interpretations of some issue areas to be more or less legitimate than others. But when states comply voluntarily with adverse CAT decisions, we may infer, in the absence of plausible alternative explanations (see similarly Deitelhoff 2009: 46), that their “normativity” likely played a role.

Recognizing the existence of nonlegal bindingness facilitates explaining the behavioral effects of CAT decisions without having to make argumentative contortions to bring together nonbindingness with an obligation/expectation of compliance. Indeed, that treaty body views can have a binding effect is implied when commentators note that they are “not per se legally binding” (Seibert-Fohr and Weniger 2019: 9) (but presumably in some other way) or comprise “quasi-legal norms that do not have a completely binding force” (Shikhelman 2019: 754) (but presumably some such force).

**Rationalist Factor: Expected Costs of Compliance and Noncompliance**

The degree of bindingness is but one factor that affects the likelihood of compliance. Whether to comply or not comply ultimately remains “a matter of state choice” (Haas 2000: 45). In making that choice, state actors commonly weigh different, partly countervailing factors, including, in addition to normative aspects, the expected material, political, and sovereignty costs of feasible alternatives.

Expected costs already play a role at the stage of commitment to human rights treaties and their ICPs (Hathaway 2007; Simmons 2009). Extant research has generated empirical evidence of hybridically motivated behavior in the context of the execution of regional court judgments, finding that states, while tending to comply, simultaneously seek to minimize compliance costs (Conant 2002; Beach 2005; von Staden 2018). In some contexts, expected costs may set a threshold for compliance. Up to a certain level, states honor their obligations and comply voluntarily, but beyond this, they seek to avoid these costs and choose noncompliance instead, unless
compliance is enforced (Downs et al. 1996). Wherever feasible, states will typically pick and choose obligations and measures whose costs they deem acceptable and ignore those viewed as too costly (Hillebrecht 2014: 49–51).

However, both compliance and noncompliance may involve costs. Aside from reputational effects, noncompliance may trigger enforcement action. Like the other core human rights treaties—and, indeed, most other international regimes—the Convention against Torture does not provide for any enforcement mechanism of its own and does not specify any sanctions to be imposed in the case of noncompliance. However, most enforcement mechanisms discussed in the human rights literature—peer pressure, naming-and-shaming, mobilization, and electoral contestation—would be available, as they do not depend on the legal status and can be applied to legally binding and nonlegally binding obligations alike (Bodansky 2015: 159; von Staden 2016). For example, Creamer and Simmons have argued, with respect to CAT’s state self-reporting mechanism, that providing information on a state’s human rights record can “stimulate the mobilization of domestic constituencies … setting the stage to demand, debate, and implement improvements over time” (2019: 1051 and 1052).

There are reasons to expect that (fear of) noncompliance/enforcement costs are, in most cases, not a relevant driver of state behavior in the case of adverse CAT decisions. First, in light of weak incentives and the sanctioners’ dilemma (Thompson 2009), interstate enforcement of human rights obligations is notoriously rare, infrequent in the case of grave, systematic human rights violations (Simmons 2009: 115) and even less frequent in the case of individual decisions. Second, in contrast to the state reporting procedure and the committee’s concluding observations, the issues at stake in individual communications often tend to be much more granular and limited to the individual applicant’s situation. As such, they are less likely to generate the same level of response as the identification of broader and more general human rights issues. Third, civil society actors usually operate with limited budgets and need to prioritize how to allocate their resources. From this perspective, it makes more sense to tackle widespread general problems rather than individual cases. While applicants may benefit from resulting administrative, policy, and legal changes, other relevant individual measures (e.g., compensation) might not be adopted. Fourth, treaty body views rarely receive much, if any, publicity and often are noticed only by specialists and the applicant’s immediate environments. There are exceptions when treaty body views make headlines in the respondent state and internationally, but generally, treaty bodies’ ICP output is barely visible to broader audiences.

4 While litigating noncompliance with treaty body views in domestic courts is normally not possible, views may influence the interpretation of legally binding provisions in related cases (cp. Kanetake 2018).
5 Relatedly, in her study of the implementation of Human Rights Committee views, Shikhelman (2019: 775) found no statistically significant effect from nongovernmental organization (NGO) density in a respondent state on that state’s compliance.
6 The Human Rights Committee’s 2018 decisions on the French niqab/burqa ban (CCPR/C/123/D/2807/2016 and CCPR/C/123/D/2747/2016) are such an exception.
All things considered, the enforcement of compliance with adverse decisions through the available mechanisms of either imposing costs or conditionally withholding incentives is, while not impossible, not very likely. It depends on a confluence of politically salient issues, sufficient visibility, and the mobilization of resources that is often lacking in the case of individual decisions. The costs of complying with an adverse decision are, by contrast, typically more certain and ascertainable and, thus, more consequential for the question of whether or not to comply.

Behavioral Expectations

Against the backdrop of these considerations, I expect the following four qualitative differences between respondent states and decisions in particular to affect compliance and noncompliance with adverse CAT decisions: regime type; the absence of the systematic use of political terror; the type of decision/violation; and the strength of civil society. These four, and the causal relationships they imply, are not the only ones that may shape compliance. Other plausibly influential factors discussed in the literature include state capacity (Anaya-Muñoz and Murdie 2021; Cole 2015), economic development, population size, and civil or international militarized conflicts (Neumayer 2005; Keith 1999), judicial independence (Powell and Staton 2009), national human rights institutions (Welch 2017), and “opinion clarity” (Staton and Romero 2019), among others. While alternative, or additional, specifications of compliance-related qualitative differences between countries and decisions are thus possible, for present purposes, the four factors discussed in the following sufficiently capture key fundamental differences between respondent states and types of decisions that can be expected to impinge on the presence of the nonlegal bindingness of CAT decisions, its compliance pull effects, and on the latter’s cost-related limitations.

Regime Type

It is a robust finding that the level of democracy positively correlates with human rights performance (see Hill and Watson 2019; Hug and Wegmann 2016: 592–594; von Stein 2015: 655–656 and fn. 6). However, democracies come in different guises and combine procedural and substantive elements usually identified to constitute, or contribute to, democratic governance to different extents. I expect that the nonlegal bindingness effect of CAT decisions is strongest with respect to liberal democracies, given that the protection of human rights is a constitutive element of their identity (Landman 2018: 49–50; Held 2006: chap. 3; Lührmann et al. 2018: 63). For a liberal democracy that affirms and upholds the function of rights as constraints and that voluntarily accepts CAT’s ICP, it should be difficult to disregard such a pronouncement subsequently because of its legal status without suffering some reputational damage to its identity as a liberal democracy. Instead, if the commitment to liberal democracy and CAT monitoring is sincere, the effective protection of rights should be valued more than the issue of the legal status of the decisions through which such rights are intended to be protected. In other words, “norm-congruence” (Sandholtz 2017) should trump formal status. I thus hypothesize the following:
H₁: Liberal democracies comply with CAT decisions more than other regime types.

**Absence of Uses of Political Terror**

A common-sense expectation is that states that do not systematically employ repression and political terror techniques—understood as violations of the right to life and physical integrity as well as of liberty and security of the person—will be more receptive to implementing adverse CAT decisions than those that use such techniques as a matter of policy. While the extensive use of political terror is not conducive to maintaining full-fledged liberal democracy, upholding human rights is only one aspect of liberal democracy, and it is imaginable that other defining aspects, such as the rule of law, remain operative while the government resorts to (some) repression to pursue a political agenda. By the same token, the absence of liberal democratic elements in other regime types need not necessarily correlate with systematic uses of political terror, and such regimes might be receptive to CAT decisions’ implications. Thus, I hypothesize that the absence of political terror favors compliance:

H₂: Regimes that do not employ political terror as state policy are more likely to comply with adverse CAT decisions than those that do.

**Type of Violation/Decision**

Different types of violations require different remedial measures that need to be adopted to comply with decisions that, in turn, generate compliance costs of different magnitudes. The costliest decisions generally are those that concern actual violations that (also) require general measures—for example changes in legislation, reforms in administrative practices, widespread training of security personnel, or practical measures, such as reforming the prison system—to address (and remove) systemic sources of repeat violations. In contrast, decisions that require remedies limited to the individual applicant are less costly. Then there are decisions that declare conditional violations that would occur if the respondent state were to proceed with its planned course of action but that still can be avoided by changing it. In the context of the Convention against Torture, Article 3’s *non-refoulement* provision regularly gives rise to such findings of “conditional” violations—violations that a respondent state can avoid by not deporting an applicant to a country where he or she may be in danger of being tortured.

Compliance with such decisions can be viewed as minimally costly not only because it typically requires measures that can be implemented easily but also because the respondent state can avoid the reputational implications of being identified as an actual violator that has exposed immigrants under its jurisdiction to (the threat of) torture. To be sure, compliance even in these cases is not entirely devoid of costs, as it may entail some (usually modest) expenditures related to providing residence permits and possibly subsistence payments as well as sovereignty costs as a result of an international expert body enjoining a state from implementing its national
authorities’ decisions in a domain usually viewed as a core part of state sovereignty. However, preventing torture should be preferable to letting it occur and providing a remedy afterward. Thus, I expect that compliance with findings of conditional violations should be particularly strong:

\[ H_3: \text{Compliance with CAT decisions that find conditional violations is more likely than compliance with CAT decisions that find actual past or ongoing violations.} \]

**Civil Society**

As noted, most enforcement means can be applied to law and soft law obligations alike, pressuring reluctant decision-makers to comply by increasing noncompliance costs. Mobilization and naming-and-shaming have been found to correlate with aggregate human rights improvements in several contexts (Dietrich and Murdie 2017; Murdie and Davis 2012; Simmons 2009; Hafner-Burton 2008), and it is reasonable to expect that such enforcement activities by civil society organizations (CSOs) and other actors can exert similarly positive effects with respect to compliance with CAT decisions. While enforcement ideally would be matched with specific decisions, this often is not feasible in practice because it is not clear how widely the net of relevant actors should be cast and because some activities occur underneath the public radar. However, as a proxy, it appears plausible to expect that such enforcement would be more likely in states with strong organized civil societies. Thus, I hypothesize as follows:

\[ H_4: \text{States with strong organized civil societies are more likely to comply with adverse CAT decisions than those with weak ones.} \]

**Empirical Analysis**

I use the method of qualitative comparative analysis (QCA) to examine the above-stated hypotheses. QCA is a set-theoretic method that is based on capturing the outcome of interest and the factors expected to explain the outcome (“conditions”) in terms of set memberships. Using Boolean algebra and formal logic, the relationship between these sets is then explored with a view to articulating statements about the causal necessity and sufficiency of (combinations of) conditions for a particular outcome. In the binary (“crisp-set”) version of QCA, each case either is (“1”), or is not (“0”), a member of a particular set (e.g., “European country”). Because conceptual boundaries are often not strictly binary and allow for some variation, QCA’s fuzzy-set variant (fsQCA) allows for gradations of set-membership to capture differences in degree, in addition to differences in kind, through the assignment of partial membership scores. The definition of the qualitative crossover point (“0.5”) between membership and nonmembership is key to fsQCA and must be determined on the basis of theoretically justifiable considerations (e.g., beginning at what level of GDP per capita is a state to be considered a “rich country”?). Cases with membership values above the crossover point are considered (more) inside the set of
interest and those with a value below it (more) outside of it (e.g., among “developed countries,” different degrees of development may be indicated by membership scores of “0.6,” “0.8,” and “1”). In QCA it is the qualitative differences as represented by membership and nonmembership that are considered to be causally relevant with respect to a particular outcome, not just any value or change along some continuous or multivalue variable, as is typically the case in most statistical, regression-based methods in which single-unit changes in the value of an independent variable are treated as having an effect on the dependent variable irrespective of where they occur along a given scale.

The data set used for the analysis comprises all 149 communications that resulted in findings of convention violations and that were decided between 1993, the year of CAT’s first adverse ICP decision,7 and the end of 2018. Compliance-related information is included up to the latest available follow-up report (from CAT’s 68th session in November 2019).8 Of the 149 communications in the data set, 124 (≈ 83%) are addressed at least once in the follow-up reports, revealing some information about their compliance status, while no such information is presently available for the decisions concerning the remaining 25 communications. Because the latter’s compliance status is uncertain, I conduct the principal analysis with the 124 observations for which official follow-up data exists and provide results for the full-set fsQCA separately (with tables in the Appendix), coding “no information” cases as “noncompliance” on the assumption that states should generally have an incentive to report implementation measures they have taken, so that if they do not report any, they most likely have not taken any.9

Calibration

With fuzzy-set QCA, in a process known as calibration, one must first determine the extent to which cases are part of the different sets/conditions of interest by assigning membership values in light of relevant theoretical knowledge and empirical information. The researcher must define what values in the raw data correspond to full membership (“1”), full nonmembership (“0”), at what level the crossover point (“0.5”) is to be placed, and how any membership values in between are to be allocated. The following provides this information for the outcome and the four conditions of interest (small caps indicate the conditions/sets in which observations can have membership, with abbreviated names used in the tables).

Outcome: COMPLIANCE

The outcome of interest is (actively pursued) compliance with adverse CAT decisions, that is, the extent to which states have remedied the violations found (or prevented them in

---

7 Halimi-Nedzibi v. Austria (CAT/C/11/D/8/1991) (1993).
8 See https://www.ohchr.org/EN/HRBodies/CAT/Pages/FUReports.aspx (accessed December 22, 2021).
9 It is possible that states occasionally fail to report remedial measures inadvertently, so that this approach to dealing with “no information” cases may overcount instances of noncompliance; disregarding these cases, though, likely undercounts noncompliance to an even greater extent.
non-refoulement cases) by providing reparation and preventing their recurrence. When the
committee assesses the adopted remedies as satisfactory, the result is coded as compliant
and fully within the compliance set (“1”).10 A committee assessment of a “partially satisfac-

tory solution” is coded as partial compliance and receives a membership value of “0.75”
(that is, it is partly within the compliance set). Instances in which CAT does not affirm at
least a partially satisfactory solution, or explicitly notes a lack of implementation, are cali-

brated as noncompliant (set membership value of “0”). The follow-up for some decisions
had been ended for reasons other than intentional state action or inaction (e.g., an applicant
departed voluntarily in a non-refoulement case); these decisions are assigned the threshold
value of “0.5”, that is, they are neither inside nor outside the condition of actively pursued
compliance.

**Condition 1: Liberal Democracy (LibDem)**

Because $H_1$ expects that it is the value orientation toward human rights and their
being accepted as constraints on authority that affect the acceptance of CAT deci-
sions as binding, I use the V-Dem Liberal Democracy Index (Coppedge et al. 2021a)
which measures liberal democracy as a combination of electoral democracy and the
protection of individual and minority rights (Coppedge et al. 2021b: 44, 49–50). A
key question with respect to continuous democracy measures concerns the numerical
values at which qualitative changes from one type of regime to another are observed.
With respect to V-Dem’s electoral democracy index, members of the V-Dem project
have suggested the mean, 0.5, as a switching point from a non-electoral to an elec-
toral democracy (Lührmann et al. 2018), whereas others have suggested 0.42 as
the threshold value because it minimizes mismatches with the dichotomous Boix-
Miller-Rosato index (Kasuya and Mori 2019).

For liberal democracy, no established threshold exists in the literature. For pre-

cent purposes, I use an approach that specifically examines the liberal component’s
contribution to the overall index and choose an index value of 0.63 as a crossover
point. Because of the aggregation formula underlying the V-Dem Liberal Democ-

cracy Index (Coppedge et al. 2021b: 44), this choice ensures that the liberal compo-
nent indicator’s value is at least above its mean. States with a Liberal Democracy
Index value equal to or above 0.81 (implying that the liberal component has a score
of 0.75 or higher) are assigned full set membership (“1”), while a score of 0.44 ($\hat{=}\$
liberal component score of 0.25 or less) establishes the “no membership” threshold.
Set membership values are calculated using the direct method of calibration (Ragin
2008: 89).

---

10 CAT’s follow-up assessments typically check whether the remedies identified in a given decision
have been successfully implemented, that is, they are case-specific and do not capture the broader effects
that CAT decisions may have irrespective of compliance or noncompliance in the individual applicants’
cases. These assessments should thus not be taken as indicators of the overall effectiveness, or ineffec-
tiveness, of CAT’s ICP with respect to triggering behavioral change as it may have an “impact beyond
compliance” (Engstrom 2019). That said, compliance or noncompliance with specific decisions is a
meaningful indicator of the extent to which states that accept CAT’s ICP are willing to provide individual
justice.
**Condition 2: (Absence of) POLITICAL TERROR (polterr)**

I use the Political Terror Scale (PTS) (Gibney et al. 2020) as an indicator of a country’s violations of core physical integrity rights. The PTS uses a five-point coding scheme that ranges from the absence of uses of political terror as government policy (level 1) to its widespread and systematic use across the entire population (level 5). Some authors have placed the crossover point from “soft” to “hard repression” between PTS levels 3 and 4 (Grauvogel and von Soest 2014: 642). Given that level 3 already indicates “extensive political imprisonment, or a recent history of such imprisonment,” that “[e]xecution or other political murders and brutality may be common” and that “[u]nlimited detention, with or without a trial, for political views is accepted” (Haschke 2020: 4), this is too permissive with respect to H2 which expects the absence of systematic uses of such practices as a key condition contributing to compliance with CAT views. Using the direct method of calibration, level 3 marks full membership (i.e., states employ political terror systematically), while level 1 is fully outside. The crossover point is at 2.1, just above level 2 (“limited amount of imprisonment for nonviolent political activity,” “exceptional” torture, “rare” political murder), which thus has greater partial membership in the condition’s complement. I calibrate using the PTS scores based on US State Department reports because they are more complete than those derived from Amnesty International’s assessments and examine the effect of the absence of POLITICAL TERROR (~polterr) on compliance.

**Condition 3: TYPE OF VIOLATION (type)**

In the present context, conditional violations concern violations of Article 3, the convention’s non-refoulement provision. Adverse CAT decisions involving only conditional non-refoulement violations receive a membership score of “1” and those concerning actual violations that already occurred or are ongoing receive a “0.” Some decisions combine conditional and realized violations; because of their conditional element, they receive a partial membership value of “0.75.”

**Condition 4: STRONG CIVIL SOCIETY (civsoc)**

As an indicator of country-level civil society strength, I use the V-Dem Civil Society Organization Participatory Environment indicator which captures both the number of CSOs and the level of popular involvement in them. It ranges from −5 (hardly any state-independent CSOs and little voluntary involvement) to +5 (many diverse CSOs and widespread active involvement in them) (Coppedge et al. 2021b: 194–195). Admittedly, this is an imperfect proxy for civil society enforcement, considering that it captures many CSOs with no interest in human rights, while for those that do address rights issues it does not measure whether they actually engage in the enforcement of CAT decisions specifically. However, the same is true of other indicators used in the literature, such as the number of non-governmental organizations (NGOs)
per capita or the number of international NGOs of which a state’s citizens are members (Shikhelman 2019: 764; Hafner-Burton and Tsutsui 2005: 1393; Neumayer 2005: 939). That said, it is reasonable to expect that the strength of civil society and the probability of social enforcement are positively correlated. As with respect to liberal democracy, no established thresholds exist in the literature to distinguish qualitative types. Once again using the direct method of calibration and given the four categories on which the indicator is based—two indicating weak civil societies and two indicating strong(er) ones—it appears justified to set the crossover point at the median (0), full membership at 5, and full nonmembership at -5.

Results

I analyze the data with the software fsQCA v. 3.1b (Ragin and Davey 2019), examining the necessity and sufficiency of conditions first with respect to the outcome of compliance and then with respect to noncompliance.

Analysis of Compliance

Compliance: Necessary Conditions A necessary condition is a condition present whenever the outcome is present, that is, it is a super-set of the outcome. As provided in Table 1, only the absence of political terror, with a consistency score of 0.904, passes the commonly suggested consistency threshold for a necessity of 0.9 (Schneider 2018: 247; Kahwati and Kane 2020: 121). Liberal democracy (0.893) comes close, while type (0.869) and civil society (0.816) are farther away. Exceeding or missing the threshold by only a small margin, liberal democracy and non-membership in political terror thus appear to be individually (near-) necessary for the outcome of compliance. Not surprisingly, the union of the three country characteristics that may be viewed as functional equivalents with respect to the latent condition of sincere commitment to human rights—liberal democracy, the complement of political terror, and civil society—is highly consistent (0.961): In most cases in which we observe compliance, respondent states are characterized by one of the three (and mostly combinations thereof).

All four conditions are empirically relevant in that compliance forms substantial subsets of them: Instances with the outcome and the condition present are fairly frequent, with coverage ranging from 62 to 82%. The relevance of necessity (RoN) indicator’s value—developed by Schneider and Wagemann to identify trivialness when

| Condition | Consistency | Coverage | Relevance of necessity |
|-----------|-------------|----------|------------------------|
| LIBDEM    | 0.894       | 0.731    | 0.640                  |
| ~POLTERR  | 0.904       | 0.707    | 0.586                  |
| TYPE      | 0.869       | 0.824    | 0.802                  |
| CIVSOC    | 0.816       | 0.627    | 0.511                  |
The Conditional Effectiveness of Soft Law: Compliance with…

The membership of cases across a condition and its complement is skewed strongly toward the former simply as a result of the empirical distribution of observations rather than being elicited by any causal processes at work (Schneider and Wagemann 2012: 234–237; calculated with Dușa 2019)—is lower than the coverage for all conditions as a result of set membership in the conditions and their complements being uneven. In contrast to consistency, no generally accepted standards exist for coverage and relevance. Schneider and Wagemann (2012: 237) view a RoN value of 0.56 as indicating trivialness, while Schneider (2019: 1116 n. 17) notes that values “lower than 0.5 are reason for concern.” Corcaci (2019: 237) proposes a best practice threshold of between 0.75 and 0.8 as separating “unambiguously relevant” necessary conditions from more trivial ones. Only type meets the latter threshold, while the highly consistent union of libdem, ~polterr, and civsoc elicits a RoN score of 0.322, indicating trivial necessity. (I address why empirically trivial necessity remains a meaningful finding in the present context in the discussion section below.)

Compliance: Sufficient Conditions Sufficiency indicates that the presence of one or more conditions leads to the outcome of interest. The truth table (Table 2) reveals the limited diversity in the set of cases generated by CAT’s ICP: Of the 16 possible combinations of conditions, six are logical remainders (not shown), and of the 10 with empirical referents, the first four account for nearly 90% of the empirical cases. For further analysis, in line with prevailing standards (Rutten 2020: 18–19), I set a frequency threshold of four which retains 95% of the cases while disregarding rows with only one or two cases, and a raw consistency threshold of at least 0.8. Only one row surpasses both thresholds. All other consistency values are significantly lower, indicating that the chosen conditions distinguish different outcomes well (Radaelli and Wagemann 2019: 283). Furthermore, the high proportional reduction

| LIBDEM | POLTERR | TYPE | CIVSOC | COMPLIANCE | Frequency | Raw consistency | PRI consistency |
|--------|---------|------|--------|------------|-----------|----------------|----------------|
| 1      | 0       | 1    | 1      | 1          | 70        | 0.845          | 0.839          |
| 1      | 0       | 0    | 1      | 0          | 15        | 0.382          | 0.292          |
| 0      | 1       | 0    | 0      | 0          | 15        | 0.169          | 0.063          |
| 0      | 1       | 0    | 1      | 0          | 10        | 0.157          | 0.064          |
| 0      | 0       | 0    | 1      | 0          | 8         | 0.358          | 0.266          |
| 0      | 0       | 0    | 0      | 0          | 2         | 0.330          | 0.195          |
| 0      | 0       | 1    | 1      | 0          | 1         | 0.814          | 0.806          |
| 0      | 1       | 1    | 1      | 0          | 1         | 0.754          | 0.745          |
| 0      | 1       | 1    | 0      | 0          | 1         | 0.709          | 0.697          |
| 1      | 1       | 0    | 1      | 0          | 0         | 0.313          | 0.094          |

The percentages of cases with membership (membership value > 0.5) in the conditions are 85% (civil society), 77% (~political terror), 69% (liberal democracy), and 59% (type).
in inconsistency (PRI) score provides evidence that the first row is not simultaneously sufficient for both compliance and its complement (Schneider and Wagemann 2012: 241–244).

Considering that the truth table includes only a single sufficient row, the conservative solution necessarily comprises a conjunction of the conditions in that row (see Table 3; the intermediate solution is identical). When loosening restrictions on assumptions for treating remainders, the parsimonious solution is reduced further, identifying type as sufficient for a compliant outcome. Parsimonious solutions are sometimes criticized for involving theoretically doubtful assumptions concerning logical remainders, the absence of empirical cases, or both, but in this case, the result is theoretically and empirically plausible. Both sufficient solutions indicate high consistency levels and large coverage values (because each solution type comprises only a single path, their raw and unique coverages are identical).

### Analysis of Noncompliance

**Noncompliance: Necessary Conditions** Individually, no condition or complement expected to coincide with noncompliance reaches satisfactory consistency levels to be deemed necessary (Table 4), nor does their union (consistency: 0.844). Thus, in contrast to compliance, noncompliance does not appear to require a specific country profile or decision type but coincides with more varying sets of conditions. However, it is notable that when noncompliance exists, a strong civil society is present more consistently than a weak one.

### Table 3 Sufficient solutions for compliance (N = 124)

| Solution type       | Solution terms              | Coverage | Consistency |
|---------------------|-----------------------------|----------|-------------|
|                     |                             | Raw      | Unique      | Raw  | PRI  |
| Conservative/intermediate | LIBDEM *~POLTERR * TYPE *CIVSOC | 0.687    | 0.687       | 0.845 | 0.839 |
| Parsimonious         | TYPE                        | 0.869    | 0.869       | 0.824 | 0.819 |

### Table 4 Analysis of necessity with respect to ~compliance (N = 124)

| Condition | Consistency | Coverage | Relevance of necessity |
|-----------|-------------|----------|------------------------|
| ~LIBDEM   | 0.593       | 0.819    | 0.920                  |
| POLTERR   | 0.538       | 0.820    | 0.930                  |
| ~TYPE     | 0.770       | 0.826    | 0.889                  |
| ~CIVSOC   | 0.400       | 0.637    | 0.876                  |
| CIVSOC    | 0.652       | 0.406    | 0.397                  |

 Springer
The Conditional Effectiveness of Soft Law: Compliance with…

Noncompliance: Sufficient Conditions The truth table for \( \neg \text{COMPLIANCE} \) (see Table 5) indicates that four of the 10 combinations of conditions with empirical referents meet the 0.8 threshold of raw consistency, two of which are disregarded for further analysis because they fall below the frequency threshold of four. The other two rows’ PRI consistency scores are adequately high to rule out common subset relations. Quine-McCluskey minimization results in two solution paths (see Table 6). The conservative/intermediate solution indicates that the joint presence of political terror, realized violations (~type), and the absence of liberal democracy (~libdem) coincides with noncompliance with a high degree of consistency, covering 46.5% of cases. The parsimonious solution highlights the political terror condition’s central role as it covers almost 54% of noncompliance cases. It should be noted that three of the four conditions expected to coincide with noncompliance are already individually sufficient for the noncompliance outcome, with raw consistency scores ranging from 0.814 to 0.860 (see Appendix, Table 7). However, their very small unique coverage scores indicate that they rarely lead to membership in \( \neg \text{COMPLIANCE} \) by themselves.

Table 5  Truth table for \( \neg \text{COMPLIANCE} \) (N = 124)

| LIBDEM | POLTERR | TYPE | CIVSOC | ~COMPLIANCE | Frequency | Raw consistency | PRI consistency |
|--------|---------|------|--------|-------------|-----------|----------------|----------------|
| 1      | 0       | 1    | 1      | 0           | 70        | 0.170          | 0.139          |
| 0      | 1       | 0    | 0      | 1           | 15        | 0.940          | 0.932          |
| 1      | 0       | 0    | 1      | 0           | 15        | 0.719          | 0.678          |
| 0      | 1       | 0    | 1      | 1           | 10        | 0.943          | 0.936          |
| 0      | 0       | 0    | 1      | 0           | 8         | 0.768          | 0.734          |
| 0      | 0       | 0    | 0      | 0           | 2         | 0.838          | 0.805          |
| 0      | 1       | 1    | 0      | 0           | 1         | 0.331          | 0.303          |
| 1      | 1       | 0    | 1      | 0           | 1         | 0.929          | 0.906          |
| 0      | 0       | 1    | 1      | 0           | 1         | 0.232          | 0.195          |
| 0      | 1       | 1    | 1      | 0           | 1         | 0.281          | 0.255          |

Table 6  Sufficient solutions for \( \neg \text{COMPLIANCE} \) (N = 124)

| Solution type      | Solution terms               | Coverage | Consistency |
|--------------------|------------------------------|----------|-------------|
|                    |                              | Raw      | Unique      | Raw  | PRI |
| Conservative/intermediate | \(~\text{LIBDEM} \times \text{POLTERR} \times \neg \text{TYPE}~) | 0.465    | 0.465       | 0.932| 0.927|
| Parsimonious        | \text{POLTERR}              | 0.538    | 0.538       | 0.820| 0.808|
Alternative Indicators and Full Data Set Analysis

I re-ran the fsQCA with different indicators to test how they would affect results. For democracy, I chose the Polity IV regime-type indicator and calibrated it in line with suggested approaches as to qualitative differences between regime types,\(^\text{12}\) with a cross-over point of 5.5, full membership at 10, and full nonmembership at −6. This is not only an indicator change but also a conceptual one as the Polity IV regime-type indicator essentially measures structural elements of democracy and not specifically liberal content.\(^\text{13}\) For political terror, I alternatively used Amnesty International-based scores which are less state-friendly in a number of cases compared with those based on US State Department reports (missing values were calculated as the average of the scores of the two nearest higher and lower years). As for civil society, V-Dem also offers a Core Civil Society Index (ranging from 0 to 1) which aggregates the participation condition I use above with an assessment of the extent of state regulation and control over/repression of CSO creation and their activities (calibrated with qualitative anchors 0, 0.5, and 1).

Notably, the alternative specifications of democracy and civil society strength exceed the consistency threshold for necessity which the former specifications did not, yielding consistency scores of 0.916 (RoN: 0.519) and 0.931 (RoN: 0.401), respectively. Some countries that scored lower with respect to liberal democracy are now among the democracies that comply with CAT views, suggesting that the liberal component may be less important than expected. As regards civil society, using the Core Civil Society Index and the direct method of calibration generates mostly higher set membership scores in comparison to those generated using the CSO Participatory Environment indicator. As a result, a few additional observations now meet the necessity test with respect to compliance, resulting in a higher consistency score. Likely in part a methodological artefact due to the different construction and spread of the two indicators, the low RoN in any event indicates trivial necessity, with 105 of the 124 observations (≈85%) having at least partial membership in this condition. The score for the political terror complement based on Amnesty International reports, by contrast, drops below the necessity threshold (0.868), considering that a few cases that were coded as more outside the political terror set than within now fall squarely within the latter (their increasing PTS scores indicating a higher incidence of political terror). The pathways identified as sufficient as a result of the minimization procedure do not change when using the alternative conditions/indicators. Only the parameters of fit for the conservative/intermediate solution are affected by that change: While its consistency changes minimally (0.838), coverage increases by about 10 percentage points (0.784) as more cases are now covered by that path.

\(^{12}\) Autocracies (ranging from −10 to −6), anocracies (from −5 to +5), and democracies (+6 to +10); see http://www.systemicpeace.org/polityproject.html. Accessed January 3, 2022.

\(^{13}\) The two indicators can be far apart. For example, in 2015, Hungary scored a perfect 10 on the Polity IV regime-type indicator but only a 0.476 on the V-Dem Liberal Democracy Index.
I also ran the original fsQCA specification with the full data set of 149 decisions, coding those 25 for which no information could be found in CAT’s follow-up reports as not being complied with (set membership value in COMPLIANCE of “0”). Analyzing compliance with the full data set, only minor changes result. Regarding necessity, consistency scores remain unaffected considering that no new compliance cases are added, but coverage falls for all conditions, and the RoN indicator slightly increases for three of them as the new cases have memberships mostly in their complements, thereby reducing skewness. With respect to sufficiency, coverage remains unchanged, while consistency falls slightly (see Appendix, Tables 8, 9, 10).

Analyzing the outcome of noncompliance (see Appendix, Tables 11, 12, 13, 14), necessity consistency rises for the conditions of interest, but the additional cases do not push any of the conditions over the 0.9 consistency threshold for necessity. The sufficiency results change as a third row now meets the consistency and frequency thresholds to be included in logical minimization. In the parsimonious solution, POLITICAL TERROR is replaced with ~LIBERAL DEMOCRACY which yields a higher consistency score and covers more cases. In the conservative/intermediate solution, in addition to the earlier path (whose consistency increases), a second path is added, combining the absence of LIBERAL DEMOCRACY with a realized violation (~TYPE) and the presence of a strong CIVIL SOCIETY. While only six cases follow this latter path (unique coverage is 4%), it indicates that some unexpected and consequential configurations exist. Individually, the sufficiency of the absence of LIBERAL DEMOCRACY, the presence of POLITICAL TERROR, and the absence of conditional violations (~TYPE) for the noncompliance outcome (~COMPLIANCE) are all strengthened by the inclusion of the additional 25 non-compliance cases (i.e., consistency increases), while the absence of a strong civil society now falls below the 0.8 threshold as a result of the cases covered by the new sufficient conjunction involving its presence.

Discussion

The fsQCA results are largely in line with expectations: When we observe compliance with adverse CAT decisions, we observe it predominantly by states that are democratic, that do not employ political terror as state policy, that feature strong civil societies, and with respect to conditional non-refoulement violations that are comparatively low-cost to implement. Viewed in the other direction, when these conditions come together and are joined with the particular decision type of conditional non-refoulement violations, the satisfactory implementation of those decisions regularly ensues; in other words, the conjunction of these conditions is sufficient for the outcome of compliance. While the analysis of an outcome (compliance) and of its complement (non-compliance) need not be mirror images of each other, here they mostly are. The absence of liberal democracy, the systematic use of political terror, and actual (as opposed to conditional) violations are strongly sufficient for the
outcome of noncompliance individually and even more so in conjunction with each other.

Civil society’s role is more Janus-faced as it appears in a conjunction sufficient for compliance and in a conjunction sufficient for noncompliance. One way to make sense of this seeming contradiction is to recognize that civil society is typically not monolithically pro-compliance but may also comprise “pro-violation constituencies” that generate political pressure against compliance with human rights norms (Cardenas 2004: 221–226; Cardenas 2007: 8–9). Such constituencies can include or be supported by the media and vested economic interests (Bernardi 2019: 233 et seq.) or may voice a dominant public opinion against the extension of rights to certain groups, such as migrants and ethnic minorities (Búzás 2018: 367–368, 374) or people in prison or preventive detention (von Staden 2018: 140, 198). When pro-violation constituencies exert greater influence over decision-makers than those that are pro-compliance, a strong civil society may thus also contribute to bringing about noncompliant outcomes.

Notably, compliance and noncompliance are largely separated along types of decisions. Decisions identifying conditional non-refoulement violations are sufficient for the outcome of compliance, whereas decisions declaring realized violations are sufficient for noncompliance. This is driven by the distribution of decision types in the data, coinciding largely with country characteristics that, likewise, favor, as the case may be, compliance or noncompliance. Complaints involving impending violations of Article 3 are predominantly lodged against liberal democracies with low PTS scores that promise to offer a safe haven, not against nondemocratic states that systematically employ torture to begin with. The latter, by contrast, mostly commit actual violations, thereby combining a lower inclination toward protecting human rights with typically costlier remedies required for compliance. Different country conditions thus lead to different types of violations, compliance with which is in turn affected by the same conditions that gave rise to these violations.

From the perspective of protecting human rights, these results are both encouraging and sobering. It is reassuring that the latent condition of sincere commitment to human rights—as reflected in liberal democracy, the absence of political terror, and a strong civil society—turns out to be sufficient for compliance with the convention’s non-refoulement norm. It was noted earlier that the necessity of the union of these three country conditions for compliance (i.e., at least one of them is observed when the compliance outcome is present) can be considered trivial (RoN: 0.338) as a result of the set being so large, with few members in its complement. Set-theoretically, given this particular distribution in the data, the assessment is correct, but in terms of substance, the result is still meaningful because it demonstrates that sincere commitment is crucial for compliance with non-refoulement cases. It also should be noted that this distribution is the result of political choices and developments across time and, thus, is different from trivial superset-subset relations that do not involve the possibility of variation and choice (such as “being a human being is trivially necessary for becoming president”). Nor is this finding diminished in importance by the fact that country characteristics rarely change quickly or profoundly for most
respondent states. While country background conditions may stay the same, each adverse decision brings into play anew the question of whether, and how, to comply in light of the specific facts of the case. The fact that countries with a sincere commitment profile generally comply with CAT non-refoulement decisions suggests that the pro-compliance effect from these conditions tends to trump any adverse case-specific considerations. In principle, many of these cases could also have gone the other way but in only a few did liberal democracies opt for noncompliance with the Committee’s decision and deport the applicants in question.\(^\text{14}\)

However, in cases of actual violations of convention rights, CAT largely fails to trigger remedial changes in most cases: Only nine out of 75 relevant decisions have been partially (six) or fully (three) complied with. No condition approaches the consistency threshold for necessity with respect to noncompliance nor does their union. Noncompliance can coincide with a more varied mix of conditions, and while liberal democracies have received a comparatively small number of non-Article 3 cases, like their nondemocratic counterparts, they are noncompliant with respect to most of them. The data indicate that once we move from conditional to actual violations, compliance deteriorates notably, even for ostensibly sincere committers. In terms of the costs of compliance, this seems to be a low threshold, but it may also reflect greater resistance triggered by the opprobrium of being charged with an actual violation, as opposed to still being able to prevent one from occurring.

While there is little smoking-gun evidence on public record acknowledging a normative compliance-pull effect of CAT decisions and the political or material boundaries to which it is subject, select country-level evidence supports the plausibility of the hybrid normative-rationalist logic here suggested. The two countries with the highest numbers of complied-with decisions (most of them concerning non-refoulement), Sweden and Switzerland,\(^\text{15}\) had both been early supporters of CAT’s ICP.\(^\text{16}\)

The Swedish Aliens Act of 2005 reflects this normative commitment by providing that residence permits should normally be granted to applicants that have successfully petitioned “an international body” in refusal-of-entry or expulsion cases.\(^\text{17}\) Switzerland, and also Denmark, systematically “reassess” or “reopen” asylum cases in response to CAT decisions, treating the latter as “new evidence” (Fox Principi 2017: 247). While these states typically comply with non-refoulement decisions by

\(^{14}\) Examples of such noncompliance are Dadar v. Canada (CAT/C/35/D/258/2004 [2005]), F. K. v. Denmark (CAT/C/56/D/580/2014 [2015]), and A.M.D. et al. v. Denmark (CAT/C/60/D/653/2015 [2017]).

\(^{15}\) In total, 32% of all adverse decisions are against Switzerland and Sweden. 22 out of 23 decisions against Sweden and 24 out of 25 decisions against Switzerland concern the issue of non-refoulement.

\(^{16}\) Sweden had included an individual complaints mechanism already in its first convention draft submitted in 1978; Switzerland later strongly supported its inclusion (Monina 2019: 585, 587), at one point expressing its preference in favor of a mandatory procedure rather than an optional one (Swiss Federal Council 1985: 296 [n. 12]).

\(^{17}\) See Swedish Aliens Act (2005), Chap. 5, Sec. 4 (“residence permits on the basis of Sweden’s international commitments”); https://www.government.se/contentassets/784b3d7be3a54a0185f284bbb2683055/aliens-act-2005_716.pdf. Accessed January 3, 2022.
adopting individual measures in the applicants’ cases, Sweden and Denmark, for example, have been resistant to changing general practices in light of the implications of repeat adverse findings, such as performing medical examinations as a matter of course in all cases of alleged past maltreatment, rather than only selectively (Scott Ford 2021: 26–28). The higher sovereignty and material costs implied by taking general measures appear to play a role here and arguably function as a scope condition for the normative compliance pull of CAT decisions. Where domestic law expressly privileges national security and anticrime considerations over non-refoulement concerns, as is the case in Canada, even the taking of individual measures may be impeded (Atak and Giffin 2018: 316, 323–325).

Does the limited effectiveness of CAT’s ICP in eliciting remedial responses make it superfluous? The answer is clearly no: Although more human rights compliance is always desirable, the effectiveness of CAT’s ICP in non-refoulement cases should not be devalued. While these decisions are comparatively easy to comply with at low material cost, they involve a core area of state sovereignty, immigration regulation. That some states regularly comply with such decisions despite the absence of a legally binding obligation to do so, notwithstanding their own domestic authorities’ prior decisions, is no small feat. In this specific area, CAT works as effectively as a court in eliciting compliance that protects fundamental physical integrity rights. What is more, for all applicants whose decisions are complied with, CAT’s decisions are clearly consequential and meaningful.

**Conclusion**

CAT’s decisions in ICP cases may be legally nonbinding, but they need not be inconsequential, frequently prompting respondent states to take measures that improve applicants’ human rights situation. I have argued that bindingness should not be viewed exclusively as a legal quality and that a duly constituted expert committee’s output can act as a normative constraint on state choice—and is, in that sense, non-legally binding—particularly for states that profess sincere commitment to human rights protections. However, there is no automatism from bindingness to compliance as competing influences are often present, including rationalist cost–benefit considerations. The empirical analysis appears to bear out the expectation of both normatively and rationally conditioned compliance. Democracies generally comply with lower-cost decisions against them, but do not perform much better than their counterparts when it comes to recognizing actual violations and providing appropriate individual and general measures to remedy them. Future research needs to examine in more detail the specific impediments to compliance in these cases as well as factors other than expected compliance costs that may explain noncompliance. Given the conditions examined here, we observe quite consistent sufficient and necessary set-relationships between their different manifestations and the outcomes of interest, confirming the expectation that sincere commitment to human rights as reflected in liberal democracy, the non-use of political terror, and a vibrant civil society matter for compliance, if only for some types of decisions. That is far from perfect but much better than no compliance at all.
### Appendix

**Table 7** Sufficiency of individual conditions for \( \neg \text{COMPLIANCE} \) (\( N = 124 \))

| Condition | Raw consistency | PRI consistency | Raw coverage | Unique coverage |
|-----------|-----------------|-----------------|--------------|----------------|
| \( \neg \text{LIBDEM} \) | 0.819 | 0.810 | 0.593 | 0.001 |
| \( \text{POLTERR} \) | 0.820 | 0.808 | 0.538 | 0.000 |
| \( \neg \text{TYPE} \) | 0.826 | 0.817 | 0.770 | 0.161 |
| \( \neg \text{CIVSOC} \) | 0.803 | 0.790 | 0.349 | 0.003 |

**Table 8** Analysis of necessity with respect to \( \text{COMPLIANCE} \) (\( N = 149 \)) (no follow-up information = noncompliance)

| Condition | Consistency | Coverage | Relevance of necessity |
|-----------|-------------|----------|------------------------|
| \( \text{LIBDEM} \) | 0.894 | 0.680 | 0.671 |
| \( \neg \text{POLTERR} \) | 0.904 | 0.647 | 0.612 |
| \( \text{TYPE} \) | 0.869 | 0.812 | 0.846 |
| \( \text{CIVSOC} \) | 0.816 | 0.543 | 0.495 |

**Table 9** Truth table for \( \text{COMPLIANCE} \) (\( N = 149 \)) (no follow-up information = noncompliance)

| LIBDEM | POLTERR | TYPE | CIVSOC | COMPLIANCE | Frequency | Raw consistency | PRI consistency |
|--------|---------|------|--------|------------|-----------|-----------------|-----------------|
| 1      | 0       | 1    | 1      | 1          | 71        | 0.833           | 0.827          |
| 0      | 1       | 0    | 0      | 0          | 25        | 0.107           | 0.038          |
| 1      | 0       | 0    | 1      | 0          | 21        | 0.287           | 0.212          |
| 0      | 1       | 0    | 1      | 0          | 16        | 0.105           | 0.041          |
| 0      | 0       | 0    | 1      | 0          | 9         | 0.293           | 0.212          |
| 0      | 0       | 0    | 0      | 0          | 3         | 0.261           | 0.148          |
| 1      | 1       | 0    | 1      | 0          | 1         | 0.280           | 0.081          |
| 0      | 0       | 1    | 1      | 0          | 1         | 0.806           | 0.797          |
| 0      | 1       | 1    | 0      | 0          | 1         | 0.702           | 0.690          |
| 0      | 1       | 1    | 1      | 0          | 1         | 0.748           | 0.739          |

**Table 10** Sufficient solutions for \( \text{COMPLIANCE} \) (\( N = 149 \)) (no follow-up information = noncompliance)

| Solution type       | Solution terms       | Coverage | Consistency |
|---------------------|----------------------|----------|-------------|
|                     |                      | Raw      | Unique      | Raw | PRI    |
| Conservative/intermediate | LIBDEM * ~ POLTERR * TYPE * CIVSOC | 0.687 | 0.687 | 0.833 | 0.827 |
| Parsimonous         | TYPE                 | 0.869    | 0.869      | 0.812 | 0.807 |
Table 11 Analysis of necessity with respect to ~COMPLIANCE (N = 149) (no follow-up information = noncompliance)

| Condition     | Consistency | Coverage | Relevance of necessity |
|---------------|-------------|----------|------------------------|
| ~LIBDEM       | 0.641       | 0.877    | 0.925                  |
| POLTERR       | 0.581       | 0.877    | 0.936                  |
| ~TYPE         | 0.829       | 0.881    | 0.891                  |
| ~CIVSOC       | 0.416       | 0.726    | 0.891                  |
| CIVSOC        | 0.620       | 0.485    | 0.465                  |

Table 12 Truth table for ~COMPLIANCE (N = 149) (no follow-up information = noncompliance)

| LIBDEM | POLTERR | TYPE | CIVSOC | ~COMPLIANCE | Frequency | Raw consistency | PRI consistency |
|--------|---------|------|--------|-------------|-----------|-----------------|-----------------|
| 1      | 0       | 1    | 1      | 0           | 71        | 0.181           | 0.151           |
| 0      | 1       | 0    | 0      | 1           | 25        | 0.962           | 0.959           |
| 1      | 0       | 0    | 1      | 0           | 21        | 0.789           | 0.766           |
| 0      | 1       | 0    | 1      | 1           | 16        | 0.961           | 0.959           |
| 0      | 0       | 0    | 1      | 1           | 9         | 0.810           | 0.788           |
| 0      | 0       | 0    | 0      | 0           | 3         | 0.872           | 0.852           |
| 1      | 1       | 0    | 1      | 0           | 1         | 0.936           | 0.919           |
| 0      | 1       | 1    | 0      | 0           | 1         | 0.338           | 0.310           |
| 0      | 1       | 1    | 1      | 0           | 1         | 0.287           | 0.261           |
| 0      | 0       | 1    | 1      | 0           | 1         | 0.239           | 0.203           |

Table 13 Sufficient solutions for ~COMPLIANCE (N = 149) (no follow-up information = noncompliance)

| Solution type       | Solution terms  | Raw coverage | Unique coverage | Consistency | Solution consistency | Solution coverage |
|---------------------|-----------------|--------------|-----------------|-------------|----------------------|-------------------|
| Conservative/intermediate | ~LIBDEM * POLTERR * ~TYPE | 0.529 | 0.214 | 0.958 | 0.930 | 0.569 |
|                     | ~LIBDEM * ~TYPE * CIVSOC | 0.355 | 0.040 | 0.917 |          |               |
| Parsimonious        | ~LIBDEM         | 0.641        | 0.641           | 0.877       | 0.877               | 0.641             |

Table 14 Sufficiency of individual conditions for ~COMPLIANCE (N = 149) (no follow-up information = noncompliance)

| Condition  | Raw consistency | PRI consistency | Raw coverage | Unique coverage |
|------------|-----------------|-----------------|--------------|----------------|
| ~LIBDEM    | 0.877           | 0.872           | 0.641        | 0.001          |
| POLTERR    | 0.877           | 0.872           | 0.581        | 0.000          |
| ~TYPE      | 0.881           | 0.877           | 0.829        | 0.148          |
| ~CIVSOC    | 0.726           | 0.707           | 0.416        | 0.030          |
Funding Open Access funding enabled and organized by Projekt DEAL. Research for this article was supported by Deutsche Forschungsgemeinschaft (DFG, German Research Foundation), project number 417704617.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

References

Abbott K. W., Keohane, R. O., Moravcsik, A., Slaughter, A.-M. & Snidal, D. (2000). The Concept of Legalization. International Organization 54:401–419.

Anaya-Muñoz, A. & Murdie, A. (2021). The Will and the Way: How State Capacity and Willingness Jointly Affect Human Rights Improvement. Human Rights Review. https://doi.org/10.1007/s12142-021-00636-y. Accessed 3 Jan 2022.

Atak, I. & Giffin, L. (2018). Canada’s Treatment of Non-Citizens through the Lens of the United Nations Individual Complaints Mechanisms. Canadian Yearbook of International Law 56: 292–327.

Beach, D. J. (2005). Why Governments Comply: An Integrative Compliance Model that Bridges the Gap between Instrumental and Normative Models of Compliance. Journal of European Public Policy 12:113–142.

Bernardi, B. B. (2019). Transitional Justice and the Inter-American Human Rights System in Peru: The Role of Anti- and Pro-Compliance Constituencies. In P. Engstrom (Ed.), The Inter-American Human System: Impact Beyond Compliance (pp. 221–246). Palgrave Macmillan, Cham, Switzerland.

Bodansky, D. (2015). Legally Binding versus Non-Legally Binding Instruments. In S. Barrett, C. Carraro & J. de Melo (Eds.), Towards a Workable and Effective Climate Regime (pp. 155–165). CEPR Press, London, and Ferdi, Paris.

Borlini, L. & Crema, L. (2019). The Legal Status of Decisions by Human Rights Treaty Bodies: Authoritative Interpretations or mission éducatrice? ESIL Conference Paper Series 13(3). https://ssrn.com/abstract=3513180. Accessed 12 Mar 2021.

Broude, T. & Shereshevsky, Y. (2021). Explaining the Practical Purchase of Soft Law: Competing and Complementary Behavioral Hypotheses. In H. G. Cohen & T. Meyer (Eds.), International Law as Behavior (pp. 98–127). Cambridge University Press, Cambridge.

Brunnée, J. & Toope, S. J. (2010). Legitimacy and Legality in International Law: An Interactional Account. Cambridge University Press, Cambridge.

Búzás, Z. I. (2018). Is the Good News About Law Compliance Good News About Norm Compliance? The Case of Racial Equality. International Organization 72:351–385.

Cardenas, S. (2004). Norm Collision: Explaining the Effects of International Human Rights Pressure on State Behavior. International Studies Review 6:213–231.

Cardenas, S. (2007). Conflict and Compliance: State Responses to International Human Rights Pressure. University of Pennsylvania Press, Philadelphia.

Cerone, J. (2016). A Taxonomy of Soft Law: Stipulating a Definition. In S. Lagoutte, T. Gammeltoft-Hansen & J. Cerone (Eds.), Tracing the Roles of Soft Law in Human Rights (pp. 15–24). Oxford University Press, Oxford.

Cole, W. (2012). Human Rights as Myth and Ceremony? Reevaluating the Effectiveness of Human Rights Treaties, 1981–2007. American Journal of Sociology 117:1131–1171.

Cole, W. (2015). Mind the Gap: State Capacity and the Implementation of Human Rights Treaties. International Organization 62:405–441.

Committee against Torture. (2002). Report of the Committee against Torture. UN Doc. A/57/44.

Committee against Torture. (2021). Report of the Committee against Torture. UN Doc. A/76/44.

Conant, L. (2002). Justice Contained: Law and Politics in the European Union. Cornell University Press, Ithaca.

Conrad, C. C. & Ritter, E. H. (2013). Treaties, Tenure, and Torture: The Conflicting Domestic Effects of International Law. The Journal of Politics 75:397–409.
Coppedge, M. et al. (2021a). V-Dem Dataset v11.1. Varieties of Democracy (V-Dem) Project. https://doi.org/10.23696/vdemds21. Accessed 28 May 2021.

Coppedge, M. et al. (2021b). V-Dem Codebook v11.1. https://doi.org/10.23696/vdemds21. Accessed 28 May 2021.

Corcaci, A. (2019). Compliance in der Europäischen Union: Mengentheoretische Konzeptformation und logische Formalisierung anhand einer QCA qualitativer Fallstudien. Springer VS, Wiesbaden.

Creamer, C. D. & Simmons, B. A. (2015). Ratification, Reporting, and Rights: Quality of Participation in the Convention against Torture. Human Rights Quarterly 37:579–608.

Creamer, C. D. & Simmons, B. A. (2019). Do Self-Reporting Regimes Matter? Evidence from the Convention against Torture. International Studies Quarterly 63:1051–1064.

Deitelhoff, N. (2009). The Discursive Process of Legalization: Charting Islands of Persuasion in the ICC Case. International Organization 63:33–65.

Dietrich, S. & Murdie, A. M. (2017). Human Rights Shaming Through INGOs and Foreign Aid Delivery. Review of International Organizations 12:95–120.

Downs, G. W., Rocke, D. M. & Barsoom, P. N. (1996). Is the Good News about Compliance Good News about Cooperation? International Organization 50:379–406.

Dușa, A. (2019). QCA with R. A Comprehensive Resource. Springer International Publishing, Cham, Switzerland.

Engstrom, P. (Ed.). (2019). The Inter-American Human System: Impact Beyond Compliance. Palgrave Macmillan, Cham, Switzerland.

Finnemore, M. & Toope, S. J. (2001). Alternatives to ‘Legalization’: Richer Views of Law and Politics. International Organization 55:743–758.

Fox Principe, Kate. (2017). Internal Mechanisms to Implement U.N. Human Rights Decisions, Notably of the U.N. Human Rights Committee. Human Rights Law Journal 37:237–250.

Franck, T. M. (1990). The Power of Legitimacy Among Nations. Oxford University Press, New York.

Gibney, M., Cornett, L., Wood, R., Haschke, P., Arnon, D., Pisanò, A., Barrett, G. & Park, B. 2020. The Political Terror Scale 1976–2019. http://www.politicalterror scale.org/. Accessed 18 Mar 2021.

Goodliffe, J. & Hawkins, D. G. (2006). Explaining Commitment: States and the Convention against Torture. Journal of Politics 68:358–371.

Grauvogel, J. & von Soest, C. (2014). Claims to Legitimacy Count: Why Sanctions Fail to Instigate Democratization in Authoritarian Regimes. European Journal of Political Research 53:635–653.

Greer, S., Gerards, J. & Slowe, R. (2018). Human Rights in the Council of Europe and the European Union: Achievements, Trends and Challenges. Cambridge University Press, Cambridge.

Guzman, A. T. & Meyer, T. L. (2010). International Soft Law. Journal of Legal Analysis 2:171–225.

Haas, P. M. (2000). Choosing to Comply: Theorizing from International Relations and Comparative Politics. In D. Shelton (Ed.), Commitment and Compliance: The Role of Non-Binding Norms in the International Legal System (pp. 43-64). Oxford University Press, Oxford.

Hafner-Burton, E. M. (2008). Sticks and Stones: Naming and Shaming the Human Rights Enforcement Problem. International Organization 62:689–716.

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

Hafner-Burton, E. M., Tsutsui, K. & Meyer, J. W. (2008). International Human Rights Law and the Politics of Legitimation: Repressive States and Human Rights Treaties. International Sociology 23:115–141.

Haschke, P. (2020). The Political Terror Scale (PTS) Codebook, Version 1.30. http://www.politicalterror scale.org/Data/Files/PTS-Codebook-V130.pdf. Accessed 18 Mar 2021.

Hathaway, O. A. (2002). Do Human Rights Treaties Make a Difference? Yale Law Journal 111:1935–2042.

Hathaway, O. A. (2007). Why Do Countries Commit to Human Rights Treaties? Journal of Conflict Resolution 51:588–621.

Hawkins, D. G. & Jacoby, W. (2010). Partial Compliance: A Comparison of the European and Inter-American Courts for Human Rights. Journal of International Law & International Relations 6:35–85.

Held, D. (2006). Models of Democracy. 3rd edn. Stanford University Press, Stanford.

Heyns, C. & Viljoen, F. (Eds.). (2002). The Impact of the United Nations Human Rights Treaties on the Domestic Level. Kluwer Law International, The Hague.

Hill, D. W. Jr. (2010). Estimating the Effects of Human Rights Treaties on State Behavior. The Journal of Politics 72:1161–1174.
The Conditional Effectiveness of Soft Law: Compliance with…

Hill, D. W. Jr. & Watson, K. A. (2019). Democracy and Compliance with Human Rights Treaties: The Conditional Effectiveness of the Convention for the Elimination of All Forms of Discrimination against Women. *International Studies Quarterly* 63:127–138.

Hillebrecht, C. (2014). *Domestic Politics and International Human Rights Tribunals: The Problem of Compliance*. Cambridge University Press, New York.

Höffinger, T. (2020). Non-binding and Therefore Irrelevant? The Global Compact for Migration. *International Journal* 75:662–673.

Hollyer, J. R. & Rosendorff, B. P. (2011). Why Do Authoritarian Regimes Sign the Convention Against Torture? Signaling, Domestic Politics, and Non-Compliance. *Quarterly Journal of Political Science* 6:275–327.

Hong, M. H. & Uzonyi, G. (2018). Deeper Commitment to Human Rights Treaties: Signaling and Investment Risk Perception. *International Interactions* 44:1040–1064.

Hug, S. & Wegmann, S. (2016). Complying with Human Rights. *International Interactions* 42:590–615.

Jurkovich, M. (2020). What Isn’t a Norm? Redefining the Conceptual Boundaries of ‘Norms’ in the Human Rights Literature. *International Studies Review* 22:693–711.

Kahwati, L. C. & Kane, H. L. (2020). Qualitative Comparative Analysis in Mixed Methods Research and Evaluation. Sage, Thousand Oaks.

Kanetake, M. (2018). UN Human Rights Treaty Monitoring Bodies Before Domestic Courts. *International & Comparative Law Quarterly* 67:201–232.

Kanetake, M. (2019). María de los Ángeles González Carreño v. Ministry of Justice, Judgment No. 1263/2018, Supreme Court of Spain, July 17, 2018. *American Journal of International Law* 113:586–592.

Karlsson-Vinkhuyzen, S. I. & Vihma, A. (2009). Comparing the Legitimacy and Effectiveness of Global Hard and Soft Law: An Analytical Framework. *Regulation & Governance* 3:400–420.

Kasuya, Y. & Mori, K. (2019). Better Regime Cutoffs for Continuous Democracy Measures. University of Gothenburg, Varieties of Democracy Institute: Users’ Working Paper No. 25.

Katzenstein, P. J. (1996). Introduction: Alternative Perspectives on National Security. In P. J. Katzenstein (Ed.), *The Culture of National Security: Norms and Identity in World Politics* (pp. 1–32). Columbia University Press, New York.

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* 36:95–118.

Kirton, J. (2019). G20 Governance for an Inclusive Liberal Order.’ *Chinese Journal of International Review* 1:1–21.

Landman, T. (2018). Democracy and Human Rights: Concepts, Measures, and Relationships. *Politics and Governance* 6:48–59.

Lührmann, A., Tannenberg, M. & Lindberg, S. I. (2018). Regimes of the World (RoW): Opening New Avenues for the Comparative Study of Political Regimes. *Politics and Governance* 6:60–77.

McQuigg, R. (2011). How Effective Is the United Nations Committee Against Torture? *European Journal of International Law* 22:813–28.

Monina, G. (2019). Article 22: Individual Complaints. In M. Nowak, M. Birk & G. Monina, (Eds.), *The United Nations Convention Against Torture and its Optional Protocol: A Commentary* (pp. 582–637). 2 ed. Oxford University Press, Oxford.

Murdie, A. M. & Davis, D. R. (2012). Shaming and Blaming: Using Events Data to Assess the Impact of Human Rights INGOs. *International Studies Quarterly* 56:1–16.

Neumayer, E. (2005). Do Human International Human Rights Treaties Improve Respect for Human Rights? *Journal of Conflict Resolution* 49:925–953.

Powell, E. J. & Staton, J. K. (2009). Domestic Judicial Institutions and Human Rights Treaty Violation. *International Studies Quarterly* 53:149–174.

Radaelli, C. M. & Wagemann, C. (2019). What Did I Leave Out? Omitted Variables in Regression and Qualitative Comparative Analysis. *European Political Science* 18:275–290.

Ragin, C. C. (2008). *Redesigning Social Inquiry: Fuzzy Sets and Beyond*. University of Chicago Press, Chicago and London.

Ragin, C. C. & Davey, S. (2019). Fuzzy-Set/Qualitative Comparative Analysis 3.1b. Irvine, California: Department of Sociology, University of California.

Raustiala, K. (2005). Form and Substance in International Agreements. *American Journal of International Law* 99:581–614.

Rottem, S. V. (2020). The Arctic Council: Between Environmental Protection and Geopolitics. Palgrave Pivot, Singapore.
Rutten, R. (2020). Applying and Assessing Large-N QCA: Causality and Robustness from a Critical Realist Perspective. Sociological Methods & Research (Online). https://doi.org/10.1177/0049124120914955. Accessed 12 March 2021.

Sandholtz, W. (2017). Domestic Law and Human Rights Treaty Commitments: The Convention against Torture. Journal of Human Rights 16:25–43.

Sartor, G. (2008). Legal Validity: An Inferential Analysis. Ratio Juris 21 (2):212–247.

Schachter, O. (1977). The Twilight Existence of Nonbinding International Agreements. American Journal of International Law 71:296–304.

Schneider, C. Q. (2018). Realists and Idealists in QCA. Political Analysis 26:246–254.

Schneider, C. Q. (2019). Two-Step QCA Revisited: The Necessity of Context Conditions. Quality & Quantity 53:1109–1126.

Schneider, C. Q. & Wagemann, C. (2012). Set-Theoretic Methods for the Social Sciences: A Guide to Qualitative Comparative Analysis. Cambridge University Press, Cambridge.

Scott Ford, S. (2021). Nordic Migration Cases before the UN Treaty Bodies: Pathways of International Accountability? iCourts Working Paper Series, no. 267. https://ssrn.com/abstract=3958974. Accessed 22 Dec 2021.

Seibert-Fohr, A. & Weniger, C. (2019). Compliance Monitoring under the International Covenant on Civil and Political Rights: To Comply, or Not to Comply, that is the Question. https://ssrn.com/abstract=3337435. Accessed 22 Dec 2021.

Shelton, D. (1997). Compliance with International Human Rights Soft Law. Studies in Transnational Legal Policy 29:119–143.

Shelton, D. (Ed). (2000). Commitment and Compliance: The Role of Non-Binding Norms in the International System. Cambridge University Press, Cambridge.

Shikhelman, V. (2019). Implementing Decisions of International Human Rights Institutions – Evidence from the United Nations Human Rights Committee. European Journal of International Law 30:753–777.

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

Staton, J. K. & Romero, A. (2019). Rational Remedies: The Role of Opinion Clarity in the Inter-American Human Rights System. International Studies Quarterly 63:477–491.

Swiss Federal Council. (1985). Botschaft betreffend das Übereinkommen gegen Folter und andere grausame, unmenschliche oder erniedrigende Behandlung oder Strafe. https://www.admin.ch/bverwaltungsbl/ viewOrigDoc.do?id=10049847. Accessed 3 Jan 2022.

Thompson, A. (2009). The Rational Enforcement of International Law: Solving the Sanctioners’ Dilemma. International Theory 1:307–321.

Ulfstein, G. (2014). Law-Making by Human Rights Treaty Bodies. In R. Liivoja & J. Petman (Eds.), International Law-Making: Essays in Honour of Jan Klabbers (pp. 249–259). Routledge, Milton Park.

United Nations Treaty Collection. (2021). Convention against Torture and Other Cruel, Inhuman or Degrading Treatment or Punishment. https://treaties.un.org/Pages/ViewDetails.aspx?REATY&mdsrg_no=IV-9&chapter=4&clang_ en. Accessed 22 Dec 2021.

van Alebeek, R. & Nolkaemper, A. (2012). The Legal Status of Decisions by Human Rights Treaty Bodies in National Law. In H. Keller & G. Ulfstein (Eds.), UN Human Rights Treaty Bodies: Law and Legitimacy (pp. 356–413). Cambridge University Press, Cambridge.

von Stein, J. (2015). Making Promises, Keeping Promises: Democracy, Ratification and Compliance in International Human Rights Law. British Journal of Political Science 46:655–679.

von Staden, A. (2016). The Political Economy of the Non-Enforcement of International Human Rights Pronouncements by States. In A. Fabbricotti (Ed.), The Political Economy of International Law: A European Perspective (pp. 230–257). Edward Elgar, Cheltenham.

von Staden, A. (2018). Strategies of Compliance with the European Court of Human Rights: Rational Choice within Normative Constraints. University of Pennsylvania Press, Philadelphia.

Vreeland, J. R. (2008). Political Institutions and Human Rights: Why Dictatorships Enter Into the United Nations Convention against Torture. International Organization 62:65–101.

Welch, R. M. (2017). National Human Rights Institutions: Domestic Implementation of International Human Rights Law. Journal of Human Rights 16:96–116.

Publisher’s Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.