Qualitative Comparative Analysis, Shaming, and International Regime Effectiveness

Olav Schram Stokke

The Fridtjof Nansen Institute, P. O. Box 326, 1326 Lysaker, Norway
E-mail: olav.s.stokke@fni.no   Tel: +47-6711 1900.

Abstract

The article presents and applies a set-theoretic comparative technique, qualitative comparative analysis (QCA), to a string of case studies on shaming as a strategy for improving the effectiveness of international regimes for resource management. This technique is particularly attractive when the number of cases available is greater than the researcher can reliably handle by narrative comparison, yet too low to support statistical procedures. QCA can capture causal conjunctions, even in small-to-intermediate-N situations, primarily because it permits the introduction of simplifying assumptions in a way that maintains a clear connection to the underlying cases – thus allowing substantive evaluation of their plausibility. A more recent fuzzy-set version lifts two limitations of the crisp-set version of QCA examined here (i.e., that variables must be dichotomous, and that the analysis makes no allowance for measurement error and non-modeled causality).

Keywords: Method; International regimes; Shaming; Comparison; Case study research.

1   Purpose and plan

Can we improve comparing case studies in order to develop general explanations of the phenomenon they examine? One approach involves narrowing in on certain causal mechanisms explanatory accounts invoke and investigating comparatively the conditions under which these causal mechanisms produce a given outcome. Ragin (1987) develops a set-theoretic comparative technique – qualitative comparative analysis (QCA) – particularly well suited for this purpose. QCA permits analytical reduction even when the number of cases is too small to support statistical analysis. The focus here is on the Boolean or crisp-set version of QCA, with some discussion of the more recent extension to fuzzy sets (Ragin, 2000; Ragin and Pennings 2005). Unlike crisp sets, fuzzy sets permit partial membership in a category – implying that both causal and outcome variables can be multichotomous.

This article examines the fruitfulness of QCA by comparing a string of case studies of efforts to solve resource management problems by means of international institutions, or regimes. Empirical studies of regime effectiveness thus far have dealt mainly with one or a handful of case studies of individual regimes at work, narratively compared (Haas, Keohane, and Levy, 1993; Stokke and Vidas, 1996; Young, 1999). Very few scholars attempt to compare the performance of numbers of regimes that would permit statistical procedures (but see Miles, Underdal, Andresen, Wettestad, Skjærseth, & Carlin, 2002). This focus on narrative comparison is in part due to the recent appearance of regime effectiveness analysis. Key concepts and theories are still in the formative stage and can be difficult to adapt and apply meaningfully outside of the issue area in which they were developed – let alone a long series of issue areas.

A frequent complaint about case-oriented research is that it often incorporates too much of the specific context of the case into the causal accounts to permit inferences to other cases: thus it tends to be weak on external validity. Others have questioned the internal validity as well, arguing that case studies are often over-determined, in the sense that the researcher is unable to choose between various possible or plausible explanations. Variable-oriented researchers call such inability a degree-of-freedom (or small-N) problem: there are too few observations relative to the set of variables under consideration to determine the relative significance of the causal factors. In an influential critique of the case-study approach, *[beginning of p. 502]* King, Keohane, and Verba (1994, p. 129) argue categorically that "nothing whatsoever can be learned about the causes of the dependent variable without taking into account other instances when the dependent variable takes on other values." For their part, many qualitative researchers maintain that the intensive approach of qualitative research allows privileged interpretive access to the complexity of the causal processes played out in any given case. Ragin (1987, p. 51) argues that the term "small-N problem" misrepresents the challenge confronting case-oriented research, since "it is not the number of relevant cases that limits the selection of method… but the nature of the method that limits the number of cases and the number of different causal conditions that the investigator is able to consider."

Use of the QCA technique will increase the number of cases and causal conditions that the investigator is able to consider. This article presents, discusses, and applies the technique to a set of case studies that involve the use of a particular regime-based strategy – shaming: that is, criticizing the target's undesirable behavior before a third party whose opinion matters to the target. The discussion here will be interspersed with contrastive side-glances to the most common variable-oriented approaches to causal inference: structured and focused comparison (George, 1979), typical of most comparative case studies of regime effectiveness to date, and statistical inference based on larger data sets.

2 QCA in a nutshell

The original, crisp-set version of QCA builds upon the binary language that George Boole developed in the mid-1800s, which also forms the mathematical basis of computer technology. Boolean algebra does not manipulate numbers but rather systematizes logical expressions in order to sort data and create a list of the configurations of circumstances associated with a given outcome. The systematic formalization inherent in QCA helps overcome two limitations of the logical analysis of causality associated with Boole's contemporary, John Stuart Mill (1904/1843). Mill's method of agreement is to search for a single common condition among cases that agree on the outcome; such conditions can be relevant to a discussion of necessity. His method of difference is to search for a single distinguishing feature among cases that disagree on the outcome – that is, experimental design (Ragin, 1987, pp. 36–42). While both may be important for discovering connections between phenomena, only the method of difference is reliable for substantiating causality. Equipped solely with the method of agreement, the investigator cannot differentiate cause from effect and will have difficulty in ruling out a spurious relationship, that is, co-variation induced by a common dependence on a non-modeled antecedent variable.

On the other hand, ascertaining that two cases, or observations, are similar in all respects save one is rarely possible outside the laboratory. Thus, in the many cases when the researcher is unable to manipulate variables, the method of difference in its direct and rigorous form is unavailable. That is why scholars often take recourse to the indirect method of difference, which involves the sequential application of the method of agreement – first on positive cases (to identify a causal candidate) and then on negative ones (to ascertain that the causal candidate is absent). This method is one of difference because the method identifies differentiating features among positive and negative cases. The method is indirect because the method proceeds, not by examining the effects of an actual change in the causal candidate, but by searching for sets of cases that agree in every respect except for the outcome and the causal candidate. To the extent that the cases actually do agree in all other relevant respects, the indirect method of difference approximates the laboratory situation and provides a natural experiment that allows the investigator to observe the effect of presence or absence of the factor under analysis. However, the number of potentially relevant factors is infinite, and so the approximation can never be complete (Mill, 1904, p. 259; Kidder, 1983, p. 38).

Unfortunately, the logical operations involved in Mill's methods of agreement and difference become increasingly complicated when the number of cases or variables grows beyond a handful. In practice, those methods do not work reliably in situations of multiple or conjunctional causation, that is, when the causes of recurrent events may lie in any of several circumstances or combinations of circumstances (Little, 1995). The set-theoretic algebra provided by the QCA technique overcomes these limitations and permits careful inspection of such configurations even with a high number of cases. This ability is particularly important in case-oriented research, which is much less inclined than variable-oriented research to address the effect of each causal condition in isolation (Ragin, 1997, pp. 36–37). Variation-oriented research also frequently addresses multiple conjunctional causations, but statistical interaction models rapidly exhaust the degrees of freedom in small- to intermediate-N situations. Even when the number of observations is high, multi-colinearity tends to make the findings indeterminate (Smith, 1990, p. 785).

The first step of a QCA analysis is to determine which set of cases to include – rarely a straightforward matter. If the research question is how international regimes can reinforce shaming efforts in international resource management, exactly what activities should count as shaming – and when can such activities be said to be regime-based? Is it enough that non-

governmental organizations criticize management practices with reference to regime standards, or should the term apply to instances when regime bodies reprimand the practices in question? Such conceptual clarification is an important part of case constitution and usually proceeds in tandem with identification of causal conditions that influence the effectiveness of shaming efforts. A back-and-forth movement occurs between the cases themselves and the concepts and theories that inform the study. That movement revolves around the building of a truth table, an ordering device that lists all possible combinations of the causal conditions that the researcher considers salient – the property space – as well as any variation in outcome among empirically observed combinations. The distribution of cases on the truth table is useful for evaluating the appropriateness of the initial model. If the model fails to align cases in reasonably comparable groupings, such heterogeneity may provide clues for *[beginning of p. 503]* refining the causal model. Such refinement is particularly important when the preliminary table includes contradictory cases – that is, when the same causal configurations yield opposite outcomes in different cases. Sometimes adding a new causal condition can remove such contradictions. Thus, the QCA procedure guides the elaboration of the causal model by confronting initial ideas with findings from the cases, thereby connecting the features of a single case with patterns across cases.

The next step is to use the truth table to evaluate propositions about necessary and sufficient conditions for the phenomenon under study. Factors that are either present or absent for *all* cases with a given outcome may, if corroborated by other knowledge, be relevant to claims about necessary conditions for that outcome. In contrast, propositions about sufficient conditions are explorable by identifying causal configurations that are present for at least one case with that outcome but in none of the cases with a different outcome. Frequency thresholds are place able on such tests requiring for instance three or more instances before a configuration is sufficient. In addition, the researcher can introduce probabilistic procedures: a causal configuration can be usually necessary or sufficient for a given outcome. Such quasi-necessity or quasi-sufficiency statements would permit one or more cases that deviate from the patterns described above – depending on the frequency distribution, the number of cases, and the probability criterion.

The initial list of configurations uniquely present in cases with a given outcome can then be minimized by means of set-theoretic algebra to reach the most general formulation of sufficient conditions compatible with the cases under study. Essentially, minimization involves removing any redundant conditions and terms found in the initial configuration.

The introduction of simplifying assumptions whenever limited diversity constrains the analysis is a final step QCA supports. Limited diversity refers to insufficient variation in the causal variables and includes empty cells in a variable-oriented data matrix or, in QCA analysis, empirical cases that fail to exhaust the number of logically possible combinations of the causal factors involved. In such situations, the researcher cannot know how the non-existing cases influence the necessity or sufficiency tests of causal configurations sketched above. Limited diversity frequently renders the analysis less conclusive. In variable-oriented analysis, this problem shows up as inability to impose statistical or comparative control on competing variables. In QCA analysis, limited diversity usually inhibits set-theoretic minimization, since few configurations will be identical in all respects save one. QCA allows substantive inspection and evaluation of the assumptions that the researcher can make about non-existing configurations in order to take the minimization further.

3   Applying QCA in empirical analysis

To illustrate the distinctive features of QCA, this article draws upon case studies of shaming in international efforts to manage marine living resources. Neither case constitution nor model specification receive much attention here, as they feature in intensive case studies reported elsewhere – on the Barents Sea cases, see Stokke, Mirovitskaya, and Anderson (1999) and Stokke (2001); on the Antarctic cases, see Stokke (1996); and on the Northwest Atlantic cases, see Stokke (2000) and Gezelius (1999). Cases revolve around situations when international regimes serve as platforms for efforts to shape problem-relevant behavior in one of the three realms of resource management: science, regulation, and compliance stimulation. Some of the cases portray different phases of a controversial issue. The rationale for viewing these phases as separate cases is that the causal configuration has shifted in that at least one of the conditions modeled in the truth table has changed its value; recall that configurative analysis conceives cases as ensembles of conditions.

3.1   Shaming in international resource management

Shaming highlights attempts to modify problem-related behavior, not by material rewards or punishment, but by *exposing* certain practices to third parties whose opinion matters to the intended target of shaming. International regimes facilitate shaming by providing, for instance, authoritative comparative standards or enhanced transparency regarding undesirable behavior.

Two regime-based conditions presumably influence the effect of such shaming: whether the shamers can substantiate their criticism by reference to explicit advice by the regime's scientific body; and whether the target behavior violates commitments assumed under the regime. Commitment here is a broad term, referring not only to legally binding conservation rules. Most fisheries management regimes provide members with either the power of veto in decisions on binding conservation measures or an opt-out clause which allows them to remain legally uncommitted even if the measure is adopted. In such situations they are nevertheless politically committed. For example, Norway and Japan have found that legal reservations on whaling cannot protect them from intensive shaming anchored in conservation measures adopted by the International Whaling Commission, of which they are both members.

In the search for non-regime conditions likely to influence the effectiveness of shaming – that is, whether or not the target adapts its behavior in response to shaming – earlier work on the subject provides several important inputs. Obviously, the target will always take into consideration the *inconvenience* of the behavioral change that shamers seek to elicit, and will balance that inconvenience against costs at two levels. Internationally, a target of shaming will typically worry about its reputation as a reliable cooperation partner, as this will affect the basis for future cooperation (Mercer, 1996). This condition is the shadow of the future: the target's perceived need to strike new deals under the regime, and whether ignoring the criticism will jeopardize such beneficial deals. The target government may also suffer domestic-level political costs if shamers scandalize it as a culprit under an international regime (Sueda and Wiseman, 1992). This suggests that shaming is more likely to succeed if there is a strong domestic constituency that can reverberate the international criticism.

These five conditional proposals are hardly daring. They reflect the pervasive expectation that political costs associating *[beginning of p. 504]*with the available options, either in relations with other states or domestically, guide governmental behavior – and that such costs are likely to be higher if authoritative scientific advice or a conservation measure

adopted by an institution that counts among its members the target of shaming can buttress shaming efforts.

3.2   The constitution of cases

The regime-based strategy under study here – shaming – is particularly effective in identifying cases: since appeals to third parties require a level of publicity or political row, the researcher is likely to discover shaming whenever it occurs. Beyond this, the constitution of cases that go into a QCA analysis proceeds in a way combining features of the two approaches that inspired it – variable- and case-oriented research. The model put up for the test should guide case selection. This rule calls for appropriate variation in the dependent variable and all the causal conditions included. Moreover, the number of cases should permit comparative or statistical control of rival conditions; and if the population so permits, the researcher should pick cases randomly to avoid systematic influence of non-modeled conditions. In contrast, case-driven narrative analyses usually start out with a few well-known cases featuring the phenomenon under scrutiny and otherwise sufficiently similar to warrant comparison. The researcher then develops the explanatory model in a step-wise manner as understanding of the cases deepens, in part by examining their similarities and dissimilarities.

The case-oriented movement back and forth between model and evidence is characteristic of QCA analysis as well. This process may add new cases to the analysis whereas others could drop out. In our context, such addition could occur because the role of shaming might have proved more marginal than initially assumed, or through narrowing down the focus to one of the versions of shaming indicated above – domestic or international. Although QCA researchers would probably be more inclined to add a causal condition deemed as substantively interesting even if such addition should lead to empty data-matrix cells, they would nevertheless heed important canons of variable-oriented research. First, received theoretical and substantive knowledge is important to the initial specification of the model, thus relating the study to a broader set of investigations of the matter (Galtung, 1969, p. 464). Second, the truth table directs attention to missing configurations, and the analytical constraints associated with limited diversity induce a search for a set of cases with adequate variance in the causal conditions. Finally, since sufficiency statements are based on causal configurations that differentiate among positive and negative outcomes, the QCA procedure concurs with the variable-oriented rule that variation in the dependent variable supports causal analysis (King et al., 1994, pp. 128–39).

Meaningful comparison also requires a level of homogeneity among the cases with regard to non-modeled conditions. This study, for instance, does not include cases scoring high on other factors likely to induce the desired behavioral change. Important examples in the context of the present study would be cases where strategies like compensation or punishment loom large. Such pragmatic case selection keeps the model as simple as possible; including cases featuring non-shaming inducements could be handled by adding conditions to the model. The price of simplicity, judged as bearable in this study, is inability to examine the effect of shaming in conjunction with such other inducements.

3.3   Three regions, ten cases

Each of the brief case sketches in this section provides the information needed to ascribe Boolean values to the five-modeled causal conditions.

3.3.1 The Barents Sea

The first instance concerns over-fishing of a Barents Sea cod stock, which is managed jointly by Norway and the Soviet Union, later Russia, since the evolution of exclusive economic zones in the late 1970s (Stokke et al., 1999). While a Joint Fisheries Commission

sets the total allowable catch because of scientific advice from the International Council for the Exploration of the Sea (ICES), annual agreements allow Norwegian coastal fishermen to continue operations with passive gear, such as hand net and long line, even after filling the quota. Coastal fishermen in Northern Norway have much appreciated this provision, which was a part of the compromise on how to divide the cod quota. From 1980 onwards, however, the abundance of fish within reach of small coastal vessels increased dramatically, pushing Norwegian over-fishing of their share of the scientifically recommended cap up to levels ranging from 40 to 100 percent. The Soviet side, unsurprisingly, expressed serious concern at Commission meetings; and the trawler-based part of the Norwegian industry, which fished on the same stock, was similarly unhappy. Mismanagement of such proportions undermined the credibility of the regime as an instrument of reciprocal restraint, implying that the shadow of the future loomed large. Then in 1983, a Protocol removed – without compensation – the right of coastal fishermen to continue fishing after the cap had been reached. The shaming had been successful. The table in section 3.4 provides the Boolean codings for this and the following nine cases.

A contrasting instance concerns regulation of the *mesh size* used in trawl bags. Citing a series of ICES recommendations for enlarging the mesh size, Norway required such enlargement of its own fishermen in the Barents Sea, and throughout the 1980s urged Soviet authorities to do the same. The main arguments for larger mesh size is protection of juvenile fish and the fact that for any given catch level, targeting bigger fish removes fewer individuals from the spawning stock. The Soviets were much less enthusiastic about this measure than the Norwegians, however, presumably in part because migration patterns make the average size of cod taken off the Norwegian coast larger than those taken off Northwest Russia. Larger mesh-size will tend to reduce the efficiency of harvesting off the Russian coast and shift the focus to the distant-water fishing grounds further west. Soviet researchers bolstered their government's position by arguing that mesh-size enlargement would *[beginning of p. 505]* increase the pressure on the mature and fertile fish, thereby depressing the stock further. Whatever the merits of such competing scientific views, the mesh-size dispute was hardly grave enough for the Soviet side to fear disruption of cooperative fisheries relations with Norway. International law ensures that states that happen to share fish stocks are stuck with each other: exclusive authority inside their respective zones implies that neither party can hope to achieve sustainable management on its own. In the early period of this dispute, there were virtually no civil society groups in the Soviet Union independent of the state apparatus to reprimand management practices if appropriate. The fisheries industry, in contrast, was well incorporated into the national and international management system, and Norway's science-based shaming efforts were rejected.

From the late 1980s, especially after the Chernobyl nuclear disaster and towards the dissolution of the Soviet Union, Moscow increasingly accepted and even encouraged domestic criticism of environmental malpractices. In this period, Soviet fisheries scientists outside of the departmental structure published various reports and articles that vehemently attacked the management approach of the Joint Fisheries Commission. By the turn of the decade, the State Committee on the Environment acquired formal competence for domestic quota allocation and fisheries compliance control (Nikitina and Pierce, 1992). Although this change added to the potential domestic reverberation of foreign criticism, the Soviet position on the mesh-size issue remained firm.

One consequence of the gradual liberalization of foreign trade in the Soviet Union, which exploded with the dissolution of that state and the emergence of Russia as legal successor in the region, was a crippling of the traditional Soviet compliance control system. This system uses a comparison of catch reports by vessels with delivery reports by processors.

From the early 1990s, Russian vessels delivered most of their harvest in Western ports, especially Norway. This change created an enforcement deficit in waters where the Norwegian Coast Guard lacked competence to inspect foreign vessels, including in the Russian zone and on the high seas. According to ICES, Russian and other vessels over-fished quotas by nearly 50 percent in 1992. The Norwegian authorities, fisheries organizations, and media sharply criticized Russia's implementation of the quota agreement. Deliveries in the West generated hard currency for the contracting Russian economy, however, so domestic reverberations were moderate. Norwegian concern grew further when Russia allowed Faroese vessels to buy parts of the already over-fished Russian quotas and operate, virtually uncontrolled, in much of the Barents Sea. As with the case of Norwegian over-fishing a decade earlier, but this time with Russia as the culprit, the entire value of the Joint Fisheries Commission as a vehicle for reciprocal cooperation was at stake. When Norway managed to substantiate that over-fishing was taking place, Russia ceded by throwing the Faroese vessels out of Russian waters and accepted a whole menu of Norwegian proposals to enhance the transparency of Russian harvesting operations.

A different challenge to the regional management system soon emerged. For decades, cod stocks had been loyal to coastal-state management aspirations and remained within national waters. Around 1990, a change in the migration pattern made cod increasingly available in the so-called *Loophole*, a sector of international waters located in-between the Norwegian and the Russian zones. Icelandic vessels, troubled by poor harvests in domestic waters, began a substantial fishery here, backed by their government, which pointed out that international law obliges coastal states to cooperate with other user states in the management of stocks that straddle national waters into the high seas (Stokke, 2001). A period of intense shaming began, in which Norway and Russia refused to negotiate the issue with Iceland. In various fora they portrayed the latter as having recklessly over-fished its own stocks and now being set to grab the fruits of coastal-state restraint in the Barents Sea. Since Iceland was not at the time incorporated in the regional regime, however, the coastal states were unable to link their criticism to Icelandic commitments to scientific or regulatory regime outputs. Indeed, the coastal-state insistence that Iceland had no rights whatsoever with regard to Barents Sea cod polarized the issue further and reduced the political clout of those in Iceland who were uneasy about their government's rapid shift from coastal-state frontrunner to distant-water hardliner (Robinson, 1996). Moreover, from the Icelandic perspective, the shadow cast by this Norwegian and Russian wrath seemed no great menace, as they had no previous fishing interests in the region to jeopardize. On the contrary, by ignoring the coastal-state complaints, which in any case were shaky on legal grounds, Iceland actually gained a level of historical fishing that could prove useful in future negotiations. The Loophole case is one of failed shaming: Iceland did sign an agreement in 1999 that acknowledged coastal-state management primacy, but only after obtaining access to cod in the Norwegian and Russian zones.

3.3.2 Northwest Atlantic

For years, Canada had faced similar problems with fishing vessels from the European Community (EC) operating on *unilateral quotas* in the Nose and Tail of the Grand Banks – a high-seas area adjacent to Canada's national waters and managed under the Northwest Atlantic Fisheries Organization (NAFO). Since 1986, the EC had taken considerable diplomatic beating from Canada for ignoring a high-seas cod moratorium advised by the NAFO Scientific Council and adopted by the regulatory Commission. Instead, the EC had set for itself quotas taken beyond Canada's jurisdiction. An opt-out clause in the NAFO Convention formally allowed such unilateral quotas, but there is little doubt that recurrent opting out of scientifically advised total quotas was contrary to the political commitment to collective management. There is no question, moreover, that this conflict severely threatened

cohesion within NAFO, with two potential costs to the target of shaming: the Canadians might – and actually did – close their own waters to EC vessels; and other NAFO members would be likely to ask why they should stay within their quotas when another major player did not. For its part, the EC portrayed NAFO as a tool for the projection of Canadian management authority and excessively cautious management principles, into high-seas areas. There is *[beginning of p. 506]* much to suggest that the EC had pinpointed the cod stock in these waters as a promising base for increasing EC catches. Such increase was vital due to the recent entry of Spain and Portugal, which had considerably enhanced EC harvesting capacity (Gezelius, 1998). Also other EC states had high stakes in this solution, because of their determination to keep large parts of the Iberian distant-water fleets outside of Community waters. Domestic reverberations of Canada's shaming were low, therefore, and the EC maintained its unilateral quota.

By 1992, however, the situation had changed in one important way: the stocks had collapsed, implying that going along with coastal-state demands would not be as inconvenient as in the past. This collapse appears to have triggered a change of heart among the parties: for several years, the EC accepted NAFO quotas and re-established cooperative ties with Canada.

3.3.3 The Antarctic

The last three cases of shaming turn on management of *krill* in the Southern Ocean. Until a decline in the activity of Eastern European states associated with their transition to market economies, the Antarctic krill had fed the largest crustacean fishery in the world. Nevertheless, parties to the 1980 Convention on the Conservation of Antarctic Marine Living Resources (CCAMLR) were slow to regulate this species. Spurred by fears that in some areas the fishery might critically interfere with the food situation for predatory species such as albatross and penguins, some members of the CCAMLR Commission had asked for restrictions. Also environmental groups criticized the lack of regulation of the largest commercial activity in the region. Until 1991, however, the Scientific Committee failed to reach agreement on the matter. When it finally did, the Commission responded by adopting the first precautionary krill cap in the area where harvesting would be most likely to affect predator species (Stokke, 1996).

The process that yielded this outcome provides interesting glimpses into the dynamic relationship between scientific investigations, shaming, and regulatory advances in Antarctic politics. The initial position of the major fishing states, especially Japan and the Soviet Union, had been that no scientific evidence indicated that the fishery posed threats to krill stocks or associated species. Accordingly, in their view, catch restrictions would not only constrain krill harvesters but also violate the principle that conservation measures should be science-based. There was rising wariness in these states that CCAMLR, like the International Whaling Commission before it, would grow into a preservationist regime (Stokke, 1991). In other words, to the extent that the user states worried about how this contested issue would impact on future cooperation, the long-term costs of accommodating shamers would seem to be higher than those associated with defiance. Both states were accomplished distant-water fishing nations, and exploitation of Antarctic resources was in no way controversial at home. The affairs of the Seventh Continent are usually rather obscure in the domestic politics of member states and are generally the purview of a thin and, in this issue area, quite autonomous bureaucratic layer. Japan and the Soviet Union rejected out of hand attempts by the conservationist majority in CCAMLR to push them into assuming commitments on the basis of possible, but unsubstantiated, ecosystem effects.

In response, the non-fishing majority in CCAMLR shifted the focus of the debate from conservation measures to information measures. More specifically, they successfully induced the Scientific Committee to request sharper and more fine-scaled reporting requirements

regarding effort and harvest. The initial reluctance on the part of the fishing nations, essentially on grounds that it would be exceedingly difficult for the fishers to maintain detailed logs of where and when, and with what result, their trawl-bags had been dragged, proved hard to defend. After all, these nations generally insisted that scientific knowledge should form the basis for harvesting regulations. Following Gorbachev's rise to power in the Soviet Union, which had rendered East–West relations in general more amiable, the cooperative atmosphere within CCAMLR had improved considerably in the mid-1980s. After the adoption of new reporting rules in 1987, the Scientific Committee was soon able to assert that krill fishery occurred in highly concentrated areas, some of which were close to important growth areas for major predators. This evidence was enough to provide the scientific basis for precautionary measures on krill.

3.4   From case studies to truth table

The first step of a truth table construction is complete: the initial specification of the model that structured the presentation of cases here. Entering the codings of those cases into the table is the next step – and frequently one that encourages refinement of the model. The ability to use contradictions as case-oriented leads for evaluating the causal model is important because QCA, like all comparative analysis, is highly vulnerable to misspecification of the model. Statements about necessary causes highlight those factors that are present or absent in all configurations with a given outcome, whereas sufficiency claims are supported by all causal configurations uniquely associated with the given outcome. The basis of both types of claims breaks down in the presence of spuriosity. On one account, the problem of missing variables is less acute in qualitative than in statistical analysis, because the intensive approach of the former ensures greater in-depth familiarity with the cases and thus less likelihood that important factors will pass ignored. On the other hand, statistical representations make allowance for misspecification by means of confidence intervals and, more generally, by the underlying assumption of probabilistic relationships among variables (Goldthorpe, 1997). The extended, fuzzy-set version of QCA is much closer to statistical approaches in this regard, as it includes a functional equivalent to confidence intervals, fuzzy adjustment, to allow for imprecise measurement of set membership as well as procedures for significance tests according to specified probability criteria (Ragin, 2000, pp. 109–16).

The number of cases reviewed above (ten) and the causal conditions modeled (five) makes undertaking systematic comparison of these cases in the same narrative mode difficult. *[beginning of p. 507]* In the truth table below, they show up as one double entry (two successful cases with the same configuration of causal condition), four successes and five failures. No contradictions occur, which may lend credibility to the model. However, the lack of contradictory cases also implies that one of the heuristic strengths of the QCA approach cannot be tried out here: its ability to support refinement of the model.

Table 1: Ten Cases of Shaming

Advice	Commitment	Shadow of the future	Inconve-nience	Rever-beration	Success	Cases
1	0	1	1	1	1	Overfishing
1	0	0	1	0	0	Mesh-size 1
1	0	0	1	1	0	Mesh-size 2
0	0	0	1	0	0	Loophole
1	1	1	1	1	1	Compliance
1	1	1	1	0	0	EC unilat. 1
1	1	1	0	0	1	EC unilat. 2
0	0	0	1	0	0	Krill cap 1
1	0	0	0	0	1	Krill report

From this table it is easy to compute an equation listing all configurations associated with a given outcome. Each such configuration is a complex statement about the absence or presence of the modeled conditions in cases with that outcome. For compactness, the equation below represents conditions by initials – upper-casing indicates presence and lower-casing absence. Mid-level dot signifies the intersection of sets and "+" signifies the union of sets.

(A) SUCCESS = A·c·S·I·R + A·C·S·I·R + A·C·S·i·r + A·c·s·i·r

Since two cases have the same causal configuration, this equation is already a step forward in descriptive compactness. Depending on the persuasiveness of the causal substantiation – the marshalling of earlier conceptual and substantive research and the construction and refinement of the truth table – it is also a statement with explanatory value. A first observation is that in all successful cases, the target of shaming had violated scientific recommendations. Accordingly, this data set is compatible with the claim that a basis in regime-generated scientific advice is a *necessary* condition for shaming to induce behavioral change. Secondly, with regard to sufficiency, the series of configurations in equation A expresses the *diversity* of combinations associated with the outcome, and supports the assumption that in social reality, there is usually more than one causal pathway to a given outcome.

3.5   Analytical reduction

The purpose of the next step is to condense equation A to expose more general patterns. This minimization occurs in three rounds. While this minimization can be done by hand even in large sets, a version of the QCA computer programs available at www.u.arizona.edu/~cragin/fsQCA/ facilitates the operation.

The first round of minimization combines features of the methods of agreement and difference by a complete series of *paired comparisons* among positive cases, allowing the removal of redundant factors within each configuration. Identifying such redundant factors is straightforward: if two terms that differ in only one causal condition yield the same outcome, the distinguishing factor must be irrelevant (Ragin, 1987, p. 93). Consider the Norwegian over-fishing of Barents Sea cod in the early 1980s and the Russian compliance failure a decade later. The two cases are equal in all (modeled) respects – except that the Russians violated an explicit regime commitment whereas the Norwegians did not. Since both are cases of successful shaming, we can conclude that whenever the shared conditions are present, the ability to base shaming in regime commitment is not essential to induce behavioral change. In other words, the causal propositions in the two initial terms are overly restrictive – and by removing the redundant condition, a more general statement replaces the two.

A caveat is in place here. While forming part of the basis of QCA, both the method of agreement and the indirect method of difference are unreliable if there is more than one sufficient cause for an outcome – a situation occurring frequently in real life and indeed the point of departure for QCA. The reason is that these methods reduce complexity by denying causal status to factors that are either absent in some positive cases (method of agreement) or present in both positive and negative cases (indirect method of difference) (Mill, 1904, pp. 255–59). Neither of those denials holds if the assumption of singular causation is relaxed,

implying that both methods can be excessively rejective of causal hypotheses. Boolean minimization ensures generalization does not occur via rejection of such a condition: the condition is rejected as causally redundant only in the specific context defined by the causal configuration in question. Accordingly, the configurative nature of QCA permits a more disciplined reduction of causal statements than do the method of agreement and the indirect method of difference.

In essence, paired comparison of configurations serves to mimic the experimental design by observing the (non)impact of an operative variable while keeping other causal candidates constant. The product of these paired comparisons is an equation listing the minimized, or least restrictive, set of causal propositions permitted by the cases examined. Since our data set *[beginning of p. 508]* has only two occurrence of cases agreeing both on the outcome and on all conditions except one, the Boolean exercise provides only a moderate reduction:

(B) SUCCESS = A·S·I·R + A·C·S·i·r + A·c·s·i·r

The next round of minimization removes terms that others contain, that is, form subsets of more general or less-restricted terms. For instance, the statement that commitment combined with non-inconvenience (C·i) is sufficient for shaming to succeed contains a statement including the same requirements and also a third – such as reverberation (C·i·r). Inspection of equation B, however, reveals that no such redundancy exists among the causal statements involved.

The same goes for the third round of simplification, relevant whenever some of the remaining terms are not essential to account for the initial configurations listed in the truth table. Such *superfluous* terms, which do not uniquely cover any original case representation, should be removed if the goal is to achieve the most parsimonious causal statement compatible with the data set. The term "superfluous" can be misleading, however, and this final round of simplification may be undesirable for other reasons as well. Whereas the first and second rounds enhance the generality of causal propositions, the third essentially narrows the range of causal pathways that the analysis lays bare. In real life, such superfluous pathways could very well be of greater practical interest than those remaining in the most compact statement – perhaps because the conditions involved in those terms prove more accessible for amendment by social actors desiring to affect the outcome.

In our data set, however, partly due to limited diversity, the second and third rounds of minimization yield no improvement in generality. The following statement rewrites equation B:

(C) SUCCESS = I·A·S·R + i· (A·c·s·r + A·C·S· r)

While the causal pattern remains rather messy, this equation clarifies that whenever the sought-for behavioral adaptation is inconvenient, shaming is unlikely to succeed unless bolstered by scientific advice as well as a high risk of cooperative breakdown and domestic reverberation. In contrast, adaptation that does not involve too much inconvenience may be induced without domestic reverberation, but only under certain conditions. There appear to be two pathways to successful shaming in such situations, the first of which is of greater substantive interest. From the first bracketed term, if the target of shaming has not committed itself to the standards that are violated, and the issue is not grave enough to threaten the regime, it would suffice that the shamer can invoke regime-based scientific advice. The

second term points to a much more demanding pathway, involving also target commitment and a risk of cooperative breakdown; and since the desired behavioral change is not too demanding or inconvenient, the success of this particular causal configuration seems over-determined.

So far, therefore, QCA analysis of the ten cases has shown a science basis to be an important factor for success in shaming efforts – this data set even suggests that it may be a necessary condition. A second observation is that violation of commitments to conservation measures – a factor often in focus among studies of regime effectiveness – does not stand out as a salient condition for shaming to succeed. One explanation may lie in the inclusion of non-binding commitments. Thirdly, equation C indicates that the more inconvenient the desired behavioral change, the more pull conditions must be present if shaming is to succeed.

3.6   Pursuing generality: assessing simplifying assumptions

With a view to reaching less restrictive causal statements, QCA allows the reasoned introduction of assumptions about the non-existing combinations of the modeled conditions. The most effective version would be to code positively (and thus add to the equation) only those non-existent combinations that would allow further simplification by pairwise comparison (Ragin, 1987, pp. 110–13). While this avenue to greater generality may seem arbitrary or even dubious, in important respects it differs favorably from the approach of statistical analysis. Many statistical packages make strong simplifying assumptions about the data with respect to homogeneity, additivity, and linearity. In contrast, a QCA researcher may or may not introduce simplifying assumptions, depending on the purpose of the investigation. Moreover, when assumptions are introduced, the researcher can *specify* them in substantive terms and thus evaluate their plausibility.

The present data involve considerable limitations on diversity, but hardly greater than in most narratively structured comparison: the limitations are expressable as 23 missing cases out of 32 logical combinations. Assuming that the non-existing configurations yield the outcome that would most *radically* minimize restrictions on the causal statement, the solution is:

(D) $\text{SUCCESS}^{\text{Minimizing}} = A \cdot i + A \cdot S \cdot R$

A procedural note is in order here. With simplifying assumptions on non-existing configurations, QCA minimization may eliminate also causal conditions that would pass the necessity test. In order to avoid such undesired exclusion, the researcher should test for necessity before sufficiency, and ensure that any necessary conditions are retained in solutions that harbor simplifying assumptions (Ragin, 2000, p. 105).

Equation D is considerably more compact and the substantive interpretation less restricted than before the introduction of simplifying assumptions. In this data set, the existence of a scientific basis stands out as a necessary condition for successful shaming; and if the desired behavioral adaptation is not inconvenient, such a basis is also sufficient. If, on the other hand, accommodation is inconvenient to the target, two additional conditions must be present: the perception that the contested issue is grave enough to threaten regime cooperation, and a strong domestic constituency that echoes the criticism.

[beginning of p. 509] QCA warns, however, that the price of achieving such a potent causal statement is to accept the plausibility of all of the following assumptions:

(E) $\text{ASSUMED SUCCESS}^{\text{Minimizing}} = A \cdot (C \cdot s \cdot i \cdot r + c \cdot S \cdot i \cdot r + c \cdot s \cdot i \cdot R + C \cdot s \cdot i \cdot R + c \cdot S \cdot i \cdot R + C \cdot s \cdot i \cdot R)$

None of those assumptions runs counter to the reasoning behind the model, or to the empirical observations of successful and failed shaming. All assumptions are loyal to the finding that a scientific basis is a necessary condition for success; none of the cases represented in the data set contradicts any of those assumptions; and none of the assumed successes involves inconvenience, the only modeled factor that makes behavioral adaptation less attractive to the target of shaming. In other data sets, however, radically minimizing assumptions may be less substantively plausible, and even here their multitude and diversity calls for some caution.

Fortunately, QCA permits more selective introduction of simplifying assumptions on non-existing configurations. For instance, the finding that regime commitment is of minor importance to the success of shaming efforts may be unconvincing. A more prudent approach would then be to remove those simplifying assumptions that harbor the claim that success is possible without commitments – that is, three of the terms in equation E. The minimized equation would then read as follows:

$$(F)\ SUCCESS^{Selective} = A \cdot (C \cdot i + s \cdot i \cdot r + S \cdot I \cdot R)$$

This more prudent set of simplifying assumptions permits causal statements that are somewhat more restricted than equation D but still rather general. As before, a scientific basis is a necessary ingredient in successful shaming. If accommodation is not inconvenient, success would in addition require either commitment, or the *absence* of two factors which are generally supportive of shaming efforts – threat to the cooperation, and domestic reverberations. Whenever the behavioral change is inconvenient or difficult, however, successful shaming would require the support of three of the four pull factors: science, the shadow of the future, and domestic reverberations.

The main point here is that as far as usage of simplifying assumptions is concerned, QCA compares favorably with both narratively structured qualitative analysis and statistical approaches. Narrative comparativists would never be able to conduct thought experiments of the type shown here with the same level of accuracy and transparency. In much statistical analysis, the assumptions imposed in order to achieve simpler representations are explicit but less readily available for substantive evaluation – partly because they are hardwired into the procedure and partly because the substantive plausibility of abstract assumptions such as homogeneity and linearity is more difficult to pin down. (Ragin, 1987).

3.7   Miscoding and robustness

Conceiving causal relationships deterministically, as the Boolean version of QCA does, makes strict assumptions as to the accuracy of the data. The threat to validity posed by sloppy or opportunistic coding is further intensified by the fact the researcher must code modeled variables dichotomously, even when cut-off points can be hard to define in terms that would settle each case unequivocally. In a strong statement of this, Goldthorpe (1997, p. 7) argues that the dichotomous-variable requirement in the QCA version used here renders the analyst unable to cope with border cases, that is, those not clearly captured by either of the categories of a dichotomous variable. Today this criticism of QCA is now largely obsolete because the fuzzy-set version permits partial membership in categories and procedures for addressing measurement error.

Even for the Boolean version, assessing the sensitivity of the findings to the coding of specific cases is a very simple matter – by recoding the conditions and rerunning the minimization. Consider the successful attempt to shame opponents of krill regulations into accepting first fine-scaled reporting and then an ecosystem-based quota in 1991. Characterizing the latter as "not inconvenient" is not entirely beyond dispute. The reasons for

this coding were that the agreed quota was higher than the preceding year's harvest and that several user-states were scaling down their operations in the region for other reasons. On the other hand, this coding would presuppose that the user-states had no worries about future implications of accepting a precautionary quota. However, the CCAMLR Commission has tended to sharpen its conservation measures over time, and the lower level of Eastern European harvesting effort might prove temporary. Recoding the cases as "inconvenient" yields no contradictions; and with the same selective assumption as above, the minimized equation would be:

(G) SUCCESS = A·(C·i + s·i·r + S·I·R)

This statement is identical with E, and the introduction of radically minimizing assumptions yields a statement equivalent with D. This equivalence means that the results are not vulnerable to the coding of the two krill cases. The more general point here is that a simple exercise permits examination of whether a result is robust to different coding of cases that may be hard to pin down on one or several causal conditions. Had such robustness been low, this would have provided a strong inducement to return to the relevant Southern Ocean cases and seek further evidence as to whether the fishing states actually perceived the krill reporting and quota measures as being inconvenient.

3.8   Negative cases and model assessment

So far, cases that feature failed efforts to pressure governments into improving their management practice by means of shaming have been peripheral to the argument here. These negative cases were indirectly used in connection with possible contradictions in the process of specifying the model, that is, by *[beginning of p. 510]* ensuring that the configurations included in the sufficiency analysis are present only in positive cases.

Making direct use of the negative cases to assess the validity of the model can be an interesting exercise. The following statement expresses the conditions, as far as our model and cases are concerned, for shaming efforts to *fail*:

(H) success = A·c·s·I·r + A·c·s·I·R +a·c·s·I·r +A·C·S·I·r +a·c·s·I·r

This equation is compatible with the hypothesis that inconvenience is a necessary condition for shaming to fail. To examine what other ingredients are necessary to guarantee failure, equation H is reducible according to the same procedures as those applied on positive cases. With radically minimizing assumptions, the result would be as follows:

(I) success = s·I + I·r

Depending on the plausibility of the assumptions underlying equation I, the equation appears to support the earlier finding that commitment is not central to the effectiveness of shaming efforts. Moreover, while a basis in authoritative scientific advice stood out as a necessary condition for successful shaming, a specific score on this factor is not a part of sufficiency statements on failure – at least not under radically minimizing assumptions. Again, a strength of the QCA procedure is that it supports substantive evaluation of the assumptions about missing configurations that permit statements as general as equations D and I.

4   Conclusions

This article has examined whether Qualitative Comparative Analysis (QCA) is a promising tool for improving comparison across detailed case studies of international regimes. The answer is affirmative because this comparative technique permits analytical reduction, or simplification, of the data without ceding on the view that comparison should proceed configuratively – even when the number of cases is fairly low.

In the *model development stage*, QCA and such variable-oriented alternatives as structured, focused comparison and statistical inference all rely upon received theoretical and substantive knowledge. If an important difference exists, it would be that the QCA procedure explicitly compels the researcher to re-examine the model in the process of constituting cases. In the Boolean version of QCA, the non-acceptance of contradictory cases – cases with the same value on the causal conditions but disagreeing on the outcome – may provide clues for model refinement that are more likely to escape non-formalized narratives and probabilistically oriented studies. On the other hand, the random element of social reality suggests that in any case such clues should be used with caution. The notion of a contradictory case is also less significant in the fuzzy-set version of QCA, since partial membership in outcome and causal categories will redefine many crisp-set contradictions to fuzzy-set discordances. Beyond this, QCA joins variable-oriented methods in encouraging strategic variation in the causal conditions specified. One difference is that QCA's ability to encompass reasoned assumptions to cope with limited diversity is likely to make users less worried about the degree-of-freedom issue and thus more prone to rest satisfied with the cases they know well and are unlikely to misrepresent. This feature is central to one of the presumed strengths of qualitative approaches.

Differences between variable-oriented techniques and QCA are greater in the stage of data analysis. The ability of the latter to capture causal conjunctions, even in small-to-intermediate-N situations, represents an important advance over statistical inference. This advantage is possible primarily because QCA permits introduction of simplifying assumptions in a way that maintains a clear connection to the underlying cases – thus allowing substantive evaluation of their plausibility. The analysis combines such justifiable counterfactual arguments with existing data in order to elicit simpler, or more general, causal accounts than those available without simplifying assumptions. At the same time, the set-theoretic algebra at the core of QCA protects this technique from the simplistic rejection of causal conditions that threatens narrative application of Mill's logical methods in the presence of multiple and conjunctional causation.

Two severe limitations of Boolean QCA – that variables must be dichotomous, and that the analysis makes no allowance for measurement error and non-modeled causality – are not present in the fuzzy-set version of QCA. So far, too few comparative studies have applied the fuzzy-set version to determine whether it will strengthen QCA's claim to the methodological middle ground between qualitative and quantitative analysis – but the potential exists. Multiple-value variables permit refining case characterization and may increase QCA analysis attractiveness to scholars inclined to in-depth narrative studies of one or a few cases. At the same time, the introduction of probability reduces the contrast between QCA and quantitative approaches to causal inference.

References

Coppedge Michael. Thickening thin concepts and theories: Combining large N and small in comparative politics. Comparative Politics 1999; 31: 465–75.

Galtung Johan. Theory and Methods of Social Research. Oslo: Universitetsforlaget, 1969.

George Alexander L. Case studies and theory development: The method of structured, focused comparison. In: Lauren PG, editor. Diplomacy: New Approaches in History, Theory, and Policy. New York: The Free Press, 1979. pp. 43–68.

Gezelius Stig S. Limits to externalisation: The EU NAFO policy 1979–97. Marine Policy 1999;23: 147–59.

[beginning of p. 511] Goldthorpe John H. Current issues in comparative macrosociology: A debate on methodological issues. Comparative Social Research 1997; 16: 1–26.

Haas Peter M Keohane Robert O Levy Marc A, editors. Institutions for the Earth: Sources of Effective International Environmental Protection. Cambridge, MA: MIT Press, 1993.

Kidder Louise H. Selltiz Wrightsman & Cooks Research Methods in Social Relations. New York: Holt, Rinehart and Winston (fourth edition; first edition 1951), 1981.

King Gary Keohane Robert O. Verba Sidney. Designing Social Inquiry: Scientific Inference in Qualitative Research. Princeton, NJ: Princeton University Press, 1994.

Little Daniel. Causal explanation in the social sciences. The Southern Journal of Philosophy 1995; 34 (Supplement): 31–56.

Mercer Jonathan. Reputation in International Politics. Ithaca, NY: Cornell University Press, 1992.

Miles Edward L Underdal Arild Andresen Steinar Wettestad Jørgen Skjærseth Jon Birger Carlin Elaine M. Environmental Regime Effectiveness: Confronting Theory with Evidence. Cambridge, MA: MIT Press, 2002.

Mill John Stuart. A System of Logic: Ratiocinative and Inductive. London: Longmans (8th edition; 1st edition 1843), 1904.

Nikitina Elena N Pearse Peter H. Conservation of marine resources in the former Soviet Union: An environmental perspective. Ocean Development and International Law 1992; 23: 369–82.

Ragin Charles C. The Comparative Method: Moving Beyond Qualitative and Quantitative Strategies. Berkeley, CA: University of California Press, 1987.

Ragin Charles C. Turning the tables: How case-oriented research challenges variable-oriented research. Comparative Social Research 1997; 16: 27–42.

Ragin, Charles C. Fuzzy-Set Social Science. Chicago, IL: University of Chicago Press, 2000.

Ragin, Charles Pennings Paul. Fuzzy sets and social research. Sociological Methods and Research 2005; 33: 423–430.

Smith Herbert L. The comparative method: Moving beyond qualitative and quantitative analysis (Ragin Charles C): Book Review. Population and Development Review 1990; 16 (4): 784–87.

Stokke Olav Schram. Transnational fishing: Japan's changing strategy. Marine Policy 1991; 15: 231–43.

Stokke Olav Schram. The effectiveness of CCAMLR. In: Stokke Olav S Vidas Davor, editors. Governing the Antarctic: The Effectiveness and Legitimacy of the Antarctic Treaty System. Cambridge: Cambridge University Press, 1996. pp. 120–51.

Stokke Olav Schram. Managing fisheries in the Barents Sea Loophole: Interplay with the UN Fish Stocks Agreement. Ocean Development and International Law 2001; 32 (3): 241–62.

Stokke Olav Schram. Managing straddling stocks: The interplay of global and regional regimes. Journal of Ocean and Coastal Management 2001; 43: 205–34.

Sueda Kiyoko Wiseman Richard L. Embarrassment remediation in Japan and the United States. International Journal of Intercultural Relations 1992; 16: 159–73.

Young Oran R, editor. The Effectiveness of International Environmental Regimes: Causal Connections and Behavioral Mechanisms. Cambridge, MA: MIT Press, 1999.