

International Regime Effectiveness

Steinar Andresen

Introduction

Unless you feel sure that international environmental regimes will have some effect, it does not make much sense to establish them. That realization motivated students of international regimes to embark on this strand of research some two decades ago. Hundreds of multilateral environmental agreements (MEAs) have been signed: indeed, in the 1990s some 20 to 30 multilateral and bilateral agreements were signed per year (Mitchell 2003: 438–439). When do these agreements solve the problems they were set out to deal with or at least contribute to a positive development – and when and why do they fail? Various large-scale projects were started in the 1990s to grapple with these questions. Through the many books and articles subsequently published we know much more about these questions today than we did some 20 years ago. Most of the research deals with “problem-solving effectiveness,” less so with questions of distribution and fairness (Young 2003). This chapter therefore focuses on problem-solving effectiveness. The main message is positive, as research confirms that “regimes do matter.” However, they do not make enough of a difference, as clearly shown in UNEP’s recent *Global Environmental Outlook* (UNEP 2012).

The bulk of the empirical research on effectiveness was conducted in the 1990s. Several major works published in the last decade, such as Miles *et al.* (2002); Breitmeyer *et al.* (2006); and Young *et al.* (2008), are also based mostly on data collected in the 1990s. To my knowledge no large-scale international research project focusing specifically on effectiveness has been launched since the turn of the millennium. This is not because this field of research is exhausted. It may be a coincidence, or it may be because attention has been diverted elsewhere, to studies on interplay between regimes and studies of partnerships. It may also reflect the relative stagnation in multilateral environmental diplomacy. There is a close connection between

the development on the environmental political arena and the research that gets conducted. Until the turn of the millennium there was rapid development and growth in the international environmental arena. This is no longer the case, as seen not least in the negotiations on the climate regime, which is where by far the most attention and resources have been invested (Victor 2011). I begin by briefly outlining the development of international environmental politics, to provide a platform for describing the subsequent development of effectiveness research. Then the focus shifts to the concept of effectiveness, what some regimes have achieved, and how this can be explained. Before concluding, there is a brief analysis of the climate regime, to shed light both on the relative standstill in multilateral environmental diplomacy and illustrate the methodological challenges.

The Development of International Environmental Politics

Although concern about nature and the natural environment arose in the nineteenth century and the first international instruments for nature protection were created at the turn of that century, it was the establishment of the United Nations that accelerated a more systematic push towards international environmental cooperation. Initially, such cooperation was quite narrow and technical, exemplified by the establishment of the UN Food and Agriculture Organization and the UN International Maritime Organization. The World Wildlife Fund, established in 1961, was the first non-governmental organization with an overtly international ambit. It marked a milestone, by introducing lobbying, campaigns, and active relations with the media. The 1960s also heralded a wider approach to environmental issues, as effects of the post-war rush to regenerate industry and manufacturing were beginning to take their toll in the form of polluted air and oceans, fanning public concern. This was, however, a typical expression of post-materialism, and was limited to the more affluent West (Andresen *et al.* 2011).

The 1972 Stockholm Conference on the Human Environment is generally seen as the watershed event that sparked a truly international approach. As a result of the Conference, the United Nations Environment Programme (UNEP) was created, important principles were adopted, and various MEAs emerged in the 1970s. Thus the Stockholm Conference scores high in terms of agenda-setting as well as institution-building (Andresen 2007a). The agreements created in the 1970s were typically *first-generation agreements*, framework conventions that essentially acknowledged the existence of a problem, without any demanding or specific obligations on the parties. The environment was fairly high on the agenda in the Western world in the 1970s, but this changed during the 1980s: the 1982 follow-up to the Stockholm Conference went virtually unnoticed by the public. And then two dramatic events changed this: the 1986 Chernobyl nuclear reactor accident, and the discovery of the hole in the ozone layer over Antarctica, attracting renewed public attention towards the environment. Both events occurred just as the UN was in the process of publishing *Our Common Future* in 1987, which introduced the concept of sustainable development. During the late 1980s, several new and more ambitious *second-generation agreements* came into being. They were more enterprising in setting numerical targets and deadlines for emission cuts. Targets were often quite random and not always well founded, but nevertheless represented a

significant advance over the first generation, because now progress or lack of such could be measured (Andresen *et al.* 2011).

With the concept of *sustainable development*, a message was sent to developing countries that the developed world was ready to accommodate calls to approach the stewardship of the environment and development as two sides of the same coin. This contributed to the inclusion of the South in the process, and sustainable development was the key focus at the 1992 Rio Summit, which became the birthplace of the Climate Convention as well the Biodiversity Convention, Agenda 21, and the Commission for Sustainable Development. As the largest international conference ever launched with very high-level political representation, the Rio Summit also marked the breakthrough for green NGO participation. Its apparent success was due not least to the widespread optimism of the early 1990s: the East–West conflict was over, economic prospects looked good, and public pressure for ambitious environmental policy was strong (Andresen 2007a).

In many ways, the 1992 Summit marked the high point of international environmental enthusiasm. Later, the media spotlight shifted focus, as did the political will and capacity to follow up the pledges made in Rio. Although the “mega-conference approach” seemed to lose some of its relevance as the follow-up was so weak, efforts to make more sophisticated MEAs continued throughout the 1990s, notably for agreements regulating air pollution, paving the way for *third-generation agreements*. It was realized that the “one-size-fits-all” approach with equal targets and timetables for all parties was not always meaningful. Cost-effectiveness and fairness were the main motivations for this more nuanced approach. Cost-effectiveness implied that cuts should be made where it would cost least, and differentiated targets thereby emerged in some regimes. The principle of fair treatment also encouraged differentiation: it was recognized as fair to give developing countries more time to reach their targets than the developed world. There was also greater focus on the need for financing and assistance to enable developing countries to meet their obligations, through such mechanisms as the Global Environment Facility.

The process of arranging mega-conferences continued with the 2002 Johannesburg Summit. This event was heavily affected by the overall international political agenda, but now in a negative way. In the aftermath of the terrorist attacks on the USA in September 2001, the times were characterized by preoccupations with war and terrorism – not the environment and development. The agenda and approach continued to broaden as poverty was now a key issue, the idea of partnerships loomed large, and the business community set their mark on the event. Despite the original intention of focusing on implementation, the summit ended up recirculating some of the ambitious aims of the UN Millennium Declaration and adding some others. The Johannesburg Summit may have had some effects in terms of discourse, but on the whole it is hard to see any substantial influence on the subsequent development of international environmental politics (Andresen 2012).

In terms of MEAs, a chemicals cluster emerged, with the Basel Convention on Hazardous Waste Management (1989) and the Rotterdam Convention on the Trade in Hazardous Chemicals (1999) supplemented by the Stockholm Convention on POPs control in 2001 (Selin 2010), and negotiations on a mercury convention are currently (May 2012) underway. However, the last decade has been characterized more by concern over too many conventions than efforts to create new ones. The system

has become incredibly complex, with demanding and often overlapping reporting requirements, particularly difficult for developing countries. There has also been concern over the conflicting commitments involved in the various regimes (Oberthür and Gehring 2006). In the run-up to the Johannesburg Summit, some policy-makers and academics called for a simplification of the system through a considerable strengthening of UNEP or even a World Environment Organization (WEO) (Biermann and Bauer 2005). The idea failed to gain momentum, and the discussion never surfaced at the Johannesburg Summit. On the other hand, this institutional architecture debate was resumed – or rather recirculated – as one of the main topics to be negotiated at the Rio+20 Conference in 2012.

In my view, the Rio+20 process as well as the global warming gridlock are further demonstrations of the recent stagnation in UN-based multilateral diplomacy. Some observers have also started to question the ability of UN multilateral diplomacy to deal effectively with the serious challenges to the environment and development, and have called for a more exclusive “club approach” (Victor 2011). Others have argued that there is no need for another global mega-conference, and the Rio+20 Conference is essentially a conference in search of a purpose (Andresen 2012).

Although the most recent period has been characterized by a rather stagnant global environmental diplomacy, tremendous progress has been achieved over a relatively short time in terms of *process and approach*: learning, knowledge, institutional design, scope of participation, as well as political effort. On this basis, the question is twofold: how and to what extent has this been translated into actual progress on the ground; and how do we measure progress or lack of such?

Regime Effectiveness: Conceptual Challenges and Solutions

The Link between the Policy Development and the Research Agendas

The focus of the International Relations (IR) community closely mirrors the political development outlined above. In the 1970s and 1980s, the establishment of international regimes was central. Why did some regimes emerge while others failed to materialize? In explaining the emergence of regimes some scholars followed the traditional IR schools of realism, liberalism, and social constructivism (Hasenclever *et al.* 1997). Others concentrated more on various types of leadership and their significance for regime creation (Young 1991). This line of research is still relevant as regimes continue to be established, although at a slower pace.

Since the early 1990s there has been a shift to studies of domestic implementation of international commitments and regime effectiveness. This reflects the fact that a significant number of MEAs have now reached maturity, having existed long enough to warrant investigation as to whether they have made a difference. Greatest attention has been paid to environmental regimes, probably because this is where the growth of regimes has been most pronounced. Some of the most important large-scale international study projects here are Brown Weiss and Jacobson (1998); Young (1999); Miles *et al.* (2002); and Breitmeier *et al.* (2006). Other significant contributions include Young (2001, 2003); Hovi *et al.* (2003a); Young *et al.* (2008); and Stokke (2012). In terms of analysis and methodology, these latter publications are very advanced, but overall they present no new empirical data on effectiveness.

In my opinion this reduces their value somewhat, as interaction between theoretical and empirical research is needed to move the field forward. In this regard, Stokke (2012) represents an interesting exception in his analysis of the effectiveness of the Barents Sea fisheries regime, combining a rigorous methodological approach with a comprehensive in-depth empirical investigation of the regime.

Most effectiveness studies have analyzed MEAs, but “soft-law” institutions have also been studied (Andresen 2007a, 2007b). As the approach has matured and developed, analysts have split regimes into different phases or particular units, in order to study changes in effectiveness over time or whether specific components were more important than others. This has enabled intra-regime comparison; comparisons between regimes and regime attributes have also been conducted (Miles *et al.* 2002). Breitmeier *et al.* (2006) have built up an International Regimes Database (IRD), in order to compare and test records on specific aspects of international regimes in relation to a range of aspects, including regime creation as well as effectiveness. Comparative quantitative analysis of these two major projects has also been conducted (Breitmeier *et al.* 2011). Another strand of effectiveness research has involved in-depth studies of domestic implementation of international regimes (Victor *et al.* 1998; Skjærseth 2000; Underdal and Hanf 2000).

Reflecting the new realities of “treaty congestion,” analysis has also been conducted on regime interplay (Oberthür and Gehring 2006; Oberthür and Stokke 2011). The initial concern was that such “congestion” might reduce the effectiveness of international environmental governance. Research has revealed a more nuanced picture of conflicts as well as synergies created through regime interplay (Skjærseth *et al.* 2006). Oberthür and Stokke (2011) have concluded that interplay management serves to enhance the effectiveness of global environmental governance.

Scholars have also dealt with partnerships between various types of actors – partnerships often established because of failure to reach effective agreements between the states concerned (Pattberg 2007; Gulbrandsen 2010; see also Chapter 23 in this volume). Within this approach, the issue of certification of forests and fisheries has loomed large. On balance, certification seems to change some management practices and create better outcomes in some cases, but it has not been regarded an effective institution for addressing some of the most serious environmental challenges. Important from our perspective is Gulbrandsen’s remark: “we still know too little about the environmental impact and efficacy of certification as a problem-solving instrument” (Gulbrandsen 2010: 180). As certification is a rather novel approach in politics and even more so in terms of academic studies, we know less about the effectiveness of this instrument compared to the more established MEAs.

Some Key Conceptual Challenges

In 1982 Oran Young wrote:

[T]here are severe limitations to what we can expect from efforts to evaluate regimes... this suggests the importance of giving some consideration to non-consequentialist approaches to the evaluation of regimes (Young 1982: 138).

As noted, this cautionary note did not prevent Young and several others from embarking on this challenge a decade later. In view of the many uncertainties and

shortcomings still existing, his warning may have been timely, but our insight, understanding, and cumulative knowledge have also grown considerably over this relatively short period.

Initially the goal of the regime was used as a measuring rod to establish the effectiveness of the regime: the higher the level of goal achievement, the higher the effectiveness of the regime in question (Andresen and Wettestad 1995). Although some analysts still apply this measure, it is problematic to use, as goals are often vague and/or overly ambitious. Consider for example the official goal of the World Health Organization, “Health for all”: not a very useful standard against which to measure progress. Moreover, goals may differ greatly in terms of ambitiousness as well as specificity, making comparison between regimes exceedingly difficult.

For quite some time there has been consensus within the “effectiveness community” on three criteria for measuring the dependent variable: *output*, *outcome*, and *impact*.

Output: Output deals with the rules and regulations reflected in the relevant regime, so this indicator is often used by legal scholars. This has been labeled level 1 implementation; level 2 output is when *formal* domestic measures are taken to comply with the international commitments (Underdal 2002a: 7). As a point of departure, the more stringent and demanding the rules regulating the behavior of the parties, the higher the effectiveness of the regime is likely to be. For example we would expect second-generation agreements to have a higher potential for being effective than first-generation agreements. Similarly, the existence of an effective compliance mechanism would seem to indicate higher effectiveness than the absence of such rules. Still, it is an open empirical question whether this potential is realized. For example, from reporting from members, it seemed that the whaling nations were all in compliance with the quotas set in the 1950s by the International Whaling Commission (IWC). Later documentation has revealed massive cheating on the part of the Soviet Union in this period. Moreover, research conducted by the IWC Scientific Committee in the 1960s showed that the quota set at the time was five times higher than it should have been (Andresen 2000). Thus, merely being in compliance can be a far cry from being effective. The output indicator deals essentially with *potential* effectiveness. Also, given the severe methodological challenges in the causal substantiation of institutional effects as to interplay, most of the evidence concerns effects at the output level, and not the outcome or impact levels (Oberthür and Stokke 2011: 318). This indicator is therefore quite weak in terms of validity; on the other hand, it is easy to measure.

Outcome: The outcome indicator seeks to measure actual progress as shown by behavioral change among key target groups in the “right” direction (regime implementation) caused by the regime. This indicator scores high in terms of validity, but involves severe obstacles in methodology. The key challenge is to establish a causal link between the regime and behavior. Consider, for instance, the massive reductions in CO₂ emissions in states with economies in transition in the 1990s. This was not a result of the effectiveness or the “bite” of the climate regime, but was due to economic recession.

Impact: The methodological challenges are even more severe when it comes to the impact indicator, the link between the regime and the effect on the problem at hand. This is the ultimate question we want to answer – the extent to which the regime has

been able to solve the problem it was set up to deal with. Unfortunately the influence of other drivers is so strong and difficult to measure that applying this indicator is usually exceedingly difficult. Consider, for example, all the many factors, apart from the regime, that are of significance for the status of given fish stock, or the level of air or ocean pollution. Consequently, this measuring rod must be used with great caution.

In short, studying outputs is usually a necessary starting point, but this needs to be supplemented with studies of outcomes to enable a better grasp on what is happening in practice. Impact indicators are so demanding in terms of methodology that they are difficult to apply in empirical studies.

We then turn to the question, and controversy, of measuring the dependent variable effectiveness. In the Miles *et al.* project, Arild Underdal (2002a) (of the “Oslo School”) used three elements to construct two different measures of effectiveness. One compared the actual performance obtained under the regime as against the no-regime counterfactual. The other compared the actual state of affairs to the best solution that could be accomplished, “the collective optimum.” The first question was used as a tool to investigate whether and to what extent regimes do matter – the relative improvement brought about by the regime. The second was designed to show whether and to what extent a particular problem is in fact solved by the regime (Hovi *et al.* 2003a). Helm and Sprinz (2000) (of the “Potsdam School”) combine these components into one comprehensive measure, through a stepwise procedure arriving at “a simple coefficient of regime effectiveness that falls into a standardized interval (0, 1)” (Hovi *et al.* 2003a: 76). It gives a very “scientific” impression to have one firm figure as a precise indicator of the effectiveness of a given regime. However important this exercise may be, the point is not mathematics or quantification but the strategies identified for determining the no-regime counterfactual and the collective optimum. In their empirical research both Miles and colleagues as well as Helm and Sprinz use a structured expert-based scoring mechanism to arrive at estimates of the non-regime counterfactual. Various measures have been used to determine the collective optimum, using game theory and other approaches. In the Miles *et al.* project, key components were application of the best external sources available and ending up with the best judgment of the case-study specialist. There are other differences and similarities between the original Oslo and Potsdam Schools, but we shall not go into these here. More basic are the key challenge and dispute: how to measure the non-regime counterfactual and the collective optimum?

This discussion was taken up by Oran Young:

How can we separate the signal of regime effects from the noise arising from the impacts of a wide range of other sources that operate simultaneously in our efforts to understand regime effectiveness? (Young 2001: 100)

This “noise” was identified as a number of driving forces including various demographic, economic, political, and technological forces that interact with each other in complex ways, producing far-reaching impacts on the problem quite apart from the dynamics of the regime. And so, “how can we determine the *proportion* of observed change in the target variables?” (Young 2001: 100; emphasis added). How precisely can causality be established? There is agreement that the specification of

the no-regime counterfactual requires a causal judgment or a method of addressing what would have happened in the absence of a particular regime. Another hotly debated topic related to arriving at counterfactuals as well as the collective optimum is the use of “expert review teams,” applied in the “Oslo School,” the “Potsdam School,” and even more extensively within the International Regimes Database (IRD) (Breitmeier *et al.* 2006).

Young (2001, 2003) raises serious doubts as to the suitability of this approach, pointing out that the different disciplines look at effectiveness through various lenses, and may thus reach divergent conclusions as to the locus of the no-regime counterfactual and thereby regime effectiveness in specific cases. Within the Miles *et al.* project the experts were essentially the case-study workers, but all authors were involved in coding all cases as a means of reducing the element of subjectivity. The process was also made easier within this project (as well as others) as the coders were mainly International Relations scholars. In the IRD, two coders were used on each case to reduce the problem of subjectivity. Still, Young’s point – that other disciplines, NGOs, or policy-makers might arrive at very different conclusions – is certainly a valid one. The use of expert coders as a main benchmark for establishing effectiveness may in fact not be very scientific, as there are too many potential flaws. More generally:

[F]rom the perspective of those having to decide whether the score on some complex dimension should be “two” or “three” on a five-point schedule, this can be a very painful process (Andresen and Wettestad 2004: 64).

Based on my experience (and that of others) from the Miles *et al.* project, scoring in terms of determining relative improvement was difficult, but still more straightforward than when it came to determining the collective optimum, a rather elusive term in relation to our cases. Hovi *et al.* (2003b) note that although there may be weaknesses in the Oslo–Potsdam solution as well as more fundamentally in effectiveness studies, that should not lead us to abandon this strand of studies. The challenge to the effectiveness community now is to test out this approach as well as other potential candidates, through new and comprehensive empirical research projects.

Most debate within the research community has concerned these issues of measurement, with less discussion on the equally important question of how to explain effectiveness. Does this mean that there is agreement on how to approach this issue? No, here we find greater disparity within the research community. According to Mitchell (2010: 172) “there is, arguably, an ‘embarrassment of riches’ of variables, each with compelling logic and empirical support but which, collectively, lack logic.” Brown Weiss and Jacobson (1998) cast the net wide, relating effectiveness characteristics to the participating countries, characteristics of the international environment, as well as characteristics of the institution and the activity involved. Victor and colleagues focused particularly on systems of implementation review (Victor *et al.* 1998), while Young and colleagues highlighted various causal pathways by which institutions influence behavior (Young 1999). Breitmeier and colleagues concluded that behavior is not only shaped by consequences and appropriateness but also by discourse, legitimacy, and habit (Breitmeier *et al.* 2006: 234–235). Despite these differences, there is wide agreement that effectiveness is influenced by both

institutional and non-institutional factors. Miles and colleagues (2002) used two main explanatory perspectives: the nature of the problem and the problem-solving capacity of the relevant institutions. The more politically and intellectually “malign” a problem, the less could be expected in terms of effectiveness. Problem-solving capacity was conceived of as a function of three variables: power, leadership, and institutional set-up.

In the following I rely heavily on the perspective of Miles and colleagues. Controlling for problem structure is important in order to be able to compare regimes. Thus it may be more of an accomplishment to achieve progress with an extremely malign problem, compared to solving an exceedingly benign one. The strength of the construction of the problem-solving capacity variable is that it includes all the major schools of IR thought. As we shall see, key elements from the realist school of thought need to be included, although this is often neglected by regime analysts.

Are Regimes Effective: When and Why?

Relatively few regimes have been analyzed, compared to the vast numbers that exist. Much is known about some favored regimes, but little or nothing about many others. Moreover, in many large-scale projects, pragmatism has often been a more important criterion for case selection than the objective principles for case selection laid down in textbooks on social science methodology. That is, the expertise and interest of those participating in the project has often been decisive for case selection. This further reduces possibilities for making general claims applicable to the wider universe of international environmental regimes (Andresen and Wettestad 2004).

In the following I present some general findings based primarily on Miles and colleagues, supplemented by other findings. Based on the scoring of the case-study workers, Underdal did the math and came up with some hard figures on the effectiveness of the 15 regimes (37 units of analysis) analyzed: “The short answer is that most of the regimes included in the sample make a positive difference but fall short of providing functionally optimal solutions” (Underdal 2002b: 435). This conclusion is in line with the findings of all the other main research projects that have been conducted. That is, there is support for the counterfactual argument that, if these regimes had not existed, things would have been worse. In more than half the cases the improvement brought about by the regime is shown to be significant. But there is also a flip side: in almost half the cases, these regimes did *not* make much of a difference, which illustrates the great variation in effectiveness.

More good news. According to Miles and colleagues, in about two out of three cases, regimes served as arenas for facilitating transnational learning and contributed to strengthening the knowledge base for policy-decisions. Moreover, we need not consider the distribution of power or the structure of the problem as given. The distribution of power may be changed through multiple maneuvers

including low-cost invasion of an existing organization (IWC), entry of new actors into the activity system itself (South Pacific Tuna), shift to another arena (radioactive waste) as well as a shift to a different regulatory approach (ship-generated oil-pollution) (Underdal 2002b: 457).

There were also instances where malign problems were dealt with rather effectively, although these were the exception. Most regimes had modest beginnings – even those that eventually achieved significant results, like the Oslo Commission on dumping in the North Sea and the ozone regime. This illustrates the positive practical effects of the more sophisticated approaches for some regimes over time. Particularly attractive are the fast-track options that have been applied in both the North Sea regime and the 1979 Convention on Long-Range Transboundary Air Pollution (LRTAP). What we find is essentially a combination of “soft law” and “hard law,” where the most enthusiastic parties went ahead with ambitious soft-law commitments, which later trickled down to the more reluctant parties (Skjærseth *et al.* 2006). Creating forums for high-level leaders to meet may also have a positive effect when the problems are not too malign. However, when problems are truly malign, as with climate change, we have seen that this approach has its limitations.

Here the flip side to these positive messages should also be noted. Even though most regimes do make a positive difference, none of those studied, not even the ozone regime, have been able to solve the problems they were set up to deal with, and they have often fallen far short of the mark. Moreover, some of the overall improvements observed have been due to “fortunate circumstances” for which the regime itself can claim little or no credit. The most fundamental intervening factor was the general growth in public demand for and governmental supply of policies for environmental protection. This was clearly the case with overall developments from the early 1970s to the turn of the millennium. It may not be equally true for the more recent period as regards the main motor in this development, the OECD region. This makes the note of caution sounded by Underdal (2002b: 457) highly relevant: “the rate of success that we have observed for environmental regimes may not be easily replicated in a stagnant or declining policy field.” The relative success achieved earlier may be hard to sustain in the current period of declining public concern over environmental issues and global economic recession.

Although malign problems can be dealt with quite effectively, this does not usually happen. The combination of high malignancy and high uncertainty is particularly lethal for problem-solving, as will be illustrated regarding the climate regime below. From these findings, power emerges as the most critical factor in dealing with strongly malign problems. Although there is little evidence that the hegemonic power approach has much significance for regime creation and effectiveness (Falkner 2005), that does not mean that power as such has no explanatory power. The shifting roles played by the USA in the ozone regime and the climate regime are illustrative. As to the negative impact of scientific uncertainty, this plagued the development of the work of the Paris Commission as well as LRTAP in the initial phases. When uncertainty was reduced, this was *one* important reason for the greater effectiveness that was achieved.

Also, although some key factors *can* be manipulated, the political engineering of institutional design of regimes tends to be a difficult exercise. For example, solid institutional capacity with an active secretariat is no guarantee of success: in most organizations, the secretariats depend more on powerful actors than vice versa. Good institutional capacity is likely to be most important when it interacts with other elements working in the same (positive) direction, like leadership and consensual science. Recent research offers new insights regarding the role of secretariats.

After studying the influence of 10 international environmental bureaucracies, mostly secretariats, Biermann and Siebenhuner (2009: 345) conclude that problem structure and what they label “people and procedures” are most important in explaining their influence, whereas institutional design makes less of a difference. Expectations as to the significance of institutional design should therefore be modest, but it is certainly not irrelevant.

The significance of various decision-making procedures is a potential candidate in this regard. The main premise in the Miles *et al.* project was the traditional one, that governance systems relying on consensus will tend to produce decisions in line with the preferences of the least ambitious member of the group. Interestingly, the Breitmeier *et al.* project reached a different conclusion: “But the evidence from the IRD does not support this conclusion. Problem improvement occurs in half the cases where regimes rely on consensus rules” (2006: 231). A further point can be added to this discussion, based on my research on the IWC (Andresen 2001). The IWC is one of the rare international organizations to practice majority voting, usually seen as a means to increase the potential effectiveness of the regime. With the IWC, however, the effect has been the opposite, as it has contributed to polarizing the already heated atmosphere in the Commission.

Finally, an observation on *exactly how much* of the change in behavior can be attributed to the regime compared to other drivers. In the studies that have been carried out, several factors are typically mentioned, of which the regime is one, but there is rarely any specification as to the proportion of change caused by the regime. Usually, the combination of these factors is said to explain the overall results achieved. For instance, Wettestad’s recent study of the acid rain regime (Wettestad 2011) finds that achievements may have been the result of better scientific understanding, an advanced regime design, changed attitudes of key players, Germany’s crucial policy turn-about, and the fall of communism, as well as the increased role of the EU. On this basis it is concluded: “It is hence clearly challenging to measure the exact impact of the multilateral mechanism on these achievements” (Wettestad 2011: 35). The more general fall-back is easier: “we are reasonably sure that emissions in Europe would have been considerably higher in such a counterfactual, no-regime situation” (Wettestad 2011: 35). This point will be elaborated further in the next section on the effectiveness of the climate regime.

The Lack of Effectiveness of the Climate Regime

Since the turn of the millennium, the environmental regime that has almost monopolized attention among policy-makers, the media, and the public has been the climate regime. By “climate regime” is here understood the political process under the UN umbrella. The main outputs have been the 1992 UN Framework Convention on Climate Change, the 1997 Kyoto Protocol, and the recent decisions taken by the Conference of the Parties (Cancun Agreement, 2010, and Durban Platform, 2011) on a future climate regime. The Framework Convention is a typical first-generation agreement with few specific or demanding commitments. Consequently it has been embraced by practically all nations of the world. Considering the short time needed to negotiate it, this convention represents a necessary first step, not least in terms of building up knowledge through the detailed reporting requirements. With the

benefit of hindsight, perhaps its most important feature is the rather clear-cut – and static – division between developed and developing countries in terms of commitments (Bodansky 1993). More recently this has hampered progress during the climate negotiations.

In terms of treaty sophistication there was a significant development from the Framework Convention to the 1997 Kyoto Protocol, which was a typical third-generation agreement with both hard targets and differentiation. Moreover, three innovative market-based mechanisms were added to facilitate implementation. Thus from negotiations started in 1991 to the adoption of the Protocol the process was quite dynamic, at least compared to the lack of achievements over the last 15 years. It took four years to reach agreement on a much-diluted Marrakesh Accord (Hovi *et al.* 2003c). Another four years went by before the Protocol entered into force (2005), and in the meantime there was not much substance to negotiate over. Since the 2007 Bali COP, negotiations were intensified in order to reach agreement on a post-Kyoto regime, but progress has been modest. Postponing key decisions has characterized the process, as shown by the 2011 Durban Platform, where yet another “roadmap” has been set up in order to have a new agreement by 2015, to be in force by 2020. The Kyoto Protocol has been prolonged, but key parties have decided to leave it. Much time has also been spent on negotiating specific mechanisms like the Green Fund, technology transfer, and REDD+, but we do not yet know what significance these will have in terms of actual emissions reductions.

What about the effect of the specific measures agreed in the Protocol? Joint implementation has not been used much, but the spread of the Clean Development Mechanism (CDM) has accelerated significantly in recent years. From a methodology perspective, this mechanism has the advantage of flowing *directly* from the regime, which removes the problem of establishing causality. However, the good news stops there, as there is uncertainty and dispute over how much emissions reductions have been caused by this mechanism, in view of such factors as carbon leakage, additionality, and transaction costs.

The challenges in terms of determining causality are more difficult as regards emissions trading, as there is no trading based directly on the climate regime. Still, various national programs have obviously been *triggered* by the Protocol, of which by far the most important is the EU ETS (on emissions trading, see also Chapter 27 in this volume). With UN negotiations deadlocked, the stricter ETS over time seems likely to be more affected by internal EU dynamics than the climate regime will be. Rather, the lack of UN progress probably contributes more to provide ammunition to those forces within the EU that oppose a strong EU climate policy. There is also considerable disagreement on the extent to which the EU ETS has contributed to emission reductions that would otherwise not have occurred. Some reductions have probably been achieved, but this is linked only partially to the climate regime.

Turning to potential behavioral effects more broadly, in terms of emission reductions compared to a business-as-usual scenario, we find that the picture is even more uncertain. On the one hand global emissions are continuing to rise sharply, so the likelihood of meeting the goal of not more than a 2 °C rise in global temperature is almost non-existent. On the other hand, the increase in emissions would probably have been even higher without the regime. The best example is provided by the many measures taken in the EU and its member-states (Skjærseth and Wettestad 2008).

The climate regime has probably also had some effect for non-Annex I countries. For example, growing concerns in China with the climate issue as well as the intensity target adopted were no doubt spurred by the climate regime. However, given the recent stalemate internationally, it seems doubtful that the recent elevation of the issue on the political agenda can be causally linked to the climate regime (Stensdal 2012). In the USA, a country not noted for federal enthusiasm for climate change policies, it has been documented that there is considerable “bottom-up” climate action among various actors (Selin and VanDeveer 2009). These processes are probably driven largely by forces other than UN negotiations. In general, from emission figures and economic growth statistics, it seems that the development of the world economy has been a more important regulator than the climate regime. Thus, effectiveness is low – but equally important is the difficulty in deciding the precise effect of the climate regime.

In terms of explaining effectiveness, the extremely malign nature of the problem goes a long way towards providing an explanation. However, the problem-solving perspective also sheds light on the poor performance of the climate regime. During this process, the most powerful actor has been the USA – most of the time as a “laggard” and a “veto-power,” blocking progress. The rising influence of China has not made progress easier. The EU has had leadership ambitions, but has been too weak to generate a sufficient number of followers to move the process forward (Andresen and Boasson 2011). In terms of institutional set-up, we do know that the massive global effort has yielded few results. The clearest example of failed multi-lateral UN diplomacy is probably the process leading up to COP 15 in Copenhagen 2009. Diplomats met during six negotiation sessions prior to COP 15 – but when the high-level segment arrived in Copenhagen there was no text on the table. In short, inefficiency and high transaction costs characterize the approach, not effective problem-solving.

Conclusion

The effectiveness community agrees that “regimes matter.” We also know that there is significant variation as to how effective regimes are, and that few if any of them have managed to solve the problems that led to their being established in the first place. Although many explanatory perspectives have been used, there is consensus that institutional as well as non-institutional factors make a difference for regime performance. There is less agreement on how precisely regime effectiveness can be measured, and efforts to arrive at exact figures as an indicator of regime effectiveness remain mired in methodological difficulties. This I have sought to illustrate empirically through a brief analysis of the climate regime. Admittedly, however, my own shortcomings in game theory and quantitative methods may represent the real bottleneck in this regard.

Empirical studies of international environmental regime effectiveness peaked towards the end of the 1990s. Important analytical advances have been made since then, but as these have not been combined with new empirical research, their significance is reduced. Some may argue that effectiveness studies are no longer timely, as the research focus has shifted to other issue areas like regime interplay, influence, the role of partnerships, or regime resilience. However relevant and intriguing these approaches may be, they are no substitute for effectiveness studies.

It is now time to launch a new wave of empirical research on the effectiveness of regimes. Is it so, as I have assumed, that there has been a general lessening in effectiveness recently – or is the picture more nuanced? In any case, new research should not be restricted to replicas and updates of previous studies. Attention should also focus on non-environmental regimes, as we know that environmental regimes may not be the only, or even the most important, regimes that actually affect the environment. The scope should also be broader than problem-solving, and the significance of regimes in terms of legitimacy as well as fairness should be discussed. So far this research has been conducted almost exclusively by Western scholars, thereby reflecting Western ways of thinking. Not least because of new geopolitical realities, scholars from the emerging economies as well as more generally from the South should be included in future research. It is important also to bring in more extensive discussions on the significance of the new geopolitical realities represented by the new emerging economies for regime effectiveness.

References

- Andresen, Steinar. 2000. "The Whaling Regime." In *Science and Politics in International Environmental Regimes: Between Integrity and Involvement*, ed. Steinar Andresen, Tora Skodvin, Arild Underdal, and Jørgen Wettestad, 35–69. Manchester: Manchester University Press.
- Andresen, Steinar. 2001. "The International Whaling Regime: 'Good' Institutions but 'Bad' Politics?" In *Toward a Sustainable Whaling Regime*, ed. Robert Friedheim, 235–269. Seattle, WA: University of Washington Press.
- Andresen, Steinar. 2007a. "The Effectiveness of UN Environmental Institutions." *International Environmental Agreements: Politics, Law and Economics*, 7(4): 317–336.
- Andresen, Steinar. 2007b. "Key Actors in UN Environmental Governance: Influence, Reform and Leadership." *International Environmental Agreements: Politics, Law and Economics*, 7(4):457–468.
- Andresen, Steinar. 2012. "Do We Need More Sustainable Development Conferences?" In *Handbook of Global Environmental Politics*, 2nd edn, ed. Peter Dauverge, 87–97. Cheltenham: Edward Elgar.
- Andresen, Steinar and Elin Lerum Boasson. 2011. "International Climate Cooperation: Clear Recommendations, Weak Commitments." In *International Environmental Agreements: An Introduction*, ed. Steinar Andresen, Elin Lerum Boasson, and Geir Hønneland, 49–67. Abingdon: Routledge.
- Andresen, Steinar, Elin Lerum Boasson, and Geir Hønneland, eds. 2011. *International Environmental Agreements: An Introduction*. Abingdon: Routledge.
- Andresen, Steinar and Jørgen Wettestad. 1995. "International Problem-Solving Effectiveness: The Oslo Project Story So Far." *International Environmental Affairs*, 7(2): 127–149.
- Andresen, Steinar and Jørgen Wettestad. 2004. "Case Studies of the Effectiveness of International Environmental Regimes." In *Regime Consequences: Methodological Challenges and Research Strategies*, ed. Arild Underdal and Oran Young, 49–70. Dordrecht: Kluwer Academic.
- Biermann, Frank and Steffen Bauer, eds. 2005. *A World Environmental Organization: Solution or Threat for International Environmental Governance*. Aldershot: Ashgate.
- Biermann, Frank and Berndt Siebenhuner, eds. 2009. *Managers of Global Change: The Influence of International Environmental Bureaucracies*. Cambridge, MA: MIT Press.
- Bodansky, Daniel. 1993. "The United Nations Framework Convention on Climate Change: A Commentary." *Yale Journal of International Law*, 18: 451–558.

- Breitmeier, Helmut, Arild Underdal, and Oran R. Young. 2011. "The Effectiveness of International Environmental Regimes: Comparing and Contrasting Findings from Quantitative Research." *International Studies Review*, 13(4): 579–605. doi:10.1111/j.1468-2486.2011.01045.x.
- Breitmeier, Helmut, Oran R. Young, and Michael Zürn. 2006. *Analyzing International Environmental Regimes: From Case Study to Database*. Cambridge, MA: MIT Press.
- Brown Weiss, Edith and Harold Jacobson, eds. 1998. *Engaging Countries: Strengthening Compliance with International Environmental Accords*. Cambridge, MA: MIT Press.
- Falkner, Robert. 2005. "American Hegemony and the Global Environment." *International Studies Review*, 7(4): 585–598. doi:10.1111/j.1468-2486.2005.00534.x.
- Gulbrandsen, Lars. 2010. *Transnational Environmental Governance: The Emergence and Effects of the Certification of Forests and Fisheries*. Cheltenham: Edward Edgar.
- Hasenclever, Andreas, Peter Mayer, and Volker Rittberger. 1997. *Theories of International Regimes*. Cambridge: Cambridge University Press.
- Helm, Carsten and Detlef Sprinz. 2000. "Measuring the Effectiveness of International Environmental Regimes." *Journal of Conflict Resolution*, 44(5): 630–652. doi:10.1177/0022002700044005004.
- Hovi, Jon, Detlef Sprinz, and Arild Underdal. 2003a. "The Oslo–Potsdam Solution to Measuring Regime Effectiveness: Critique, Response, and the Road Ahead." *Global Environmental Politics*, 3(3):74–96. doi:10.1162/152638003322469286.
- Hovi, Jon, Detlef Sprinz, and Arild Underdal. 2003b. "Regime Effectiveness and the Oslo–Potsdam Solution: A Rejoinder to Oran Young." *Global Environmental Politics*, 3(3): 105–107. doi:10.1162/152638003322469303.
- Hovi, Jon, Tora Skodvin, and Steinar Andresen. 2003c. "The Persistence of the Kyoto Protocol: Why Other Annex I Countries Move On without the United States." *Global Environmental Politics*, 3(4): 1–23. doi:10.1162/152638003322757907.
- Miles, Edward L., Arild Underdal, Steinar Andresen *et al.* 2002. *Environmental Regime Effectiveness: Confronting Theory with Evidence*. Cambridge, MA: MIT Press.
- Mitchell, Ronald B. 2003. "International Environmental Agreements: A Survey of Their Features, Formation and Effects." *Annual Review of Environment and Resources*, 28, November: 429–461.
- Mitchell, Ronald B. 2010. *International Politics and the Environment*. Thousand Oaks, CA: Sage.
- Oberthür, Sebastian and Thomas Gehring, eds. 2006. *Institutional Interaction in Global Environmental Governance: Synergy and Conflict among International and EU Policies*. Cambridge, MA: MIT Press.
- Oberthür, Sebastian and Olav Schram Stokke, eds. 2011. *Managing Institutional Complexity: Regime Interplay and Global Environmental Change*. Cambridge, MA: MIT Press.
- Pattberg, Phillip. 2007. *Private Institutions and Global Governance: The New Politics of Environmental Sustainability*. Cheltenham: Edward Elgar.
- Selin, Henrik. 2010. *Global Governance of Hazardous Chemicals: Challenges of Multilevel Management*. Cambridge, MA: MIT Press.
- Selin, Henrik and Stacy D. VanDeveer. 2009. *Changing Climates in North American Politics: Institutions, Policymaking, and Multilevel Governance*. Cambridge, MA: MIT Press.
- Skjærseth, Jon Birger. 2000. *North Sea Cooperation: Linking International and Domestic Pollution Control*. Manchester: Manchester University Press.
- Skjærseth, Jon Birger, Olav Schram Stokke, and Jørgen Wettestad. 2006. "Soft Law, Hard Law, and Effective Implementation of International Environmental Norms." *Global Environmental Politics*, 6(3): 104–120. doi:10.1162/glep.2006.6.3.104.
- Skjærseth, Jon Birger and Jørgen Wettestad. 2008. *EU Emission Trading: Initiation, Decision-Making and Implementation*. Aldershot: Ashgate.
- Stensdal, Iselin. 2012. *China's Climate Policy 1998–2011: From Zero to Hero?* FNI Report 9/2012. Lysaker, Norway: Fridtjof Nansen Institute.

- Stokke, Olav Schram. 2012. *Disaggregating International Regimes. A New Approach to Evaluation and Comparison*. Cambridge, MA: MIT Press.
- Underdal, Arild. 2002a. "One Question Two Answers." In *Environmental Regime Effectiveness: Confronting Theory with Evidence*, ed. Edward L. Miles, Arild Underdal, Steinar Andresen *et al.*, 3–46. Cambridge, MA: MIT Press.
- Underdal, Arild. 2002b. "Conclusions: Patterns of Regime Effectiveness." In *Environmental Regime Effectiveness: Confronting Theory with Evidence*, ed. Edward L. Miles, Arild Underdal, Steinar Andresen *et al.*, 433–466. Cambridge, MA: MIT Press.
- Underdal, Arild and Kenneth Hanf, eds. 2000. *International Environmental Agreements and Domestic Politics: The Case of Acid Rain*. Aldershot: Ashgate.
- UNEP. 2012. *GEO 5: Global Environment Outlook*. Nairobi: United Nations Environment Programme.
- Victor, David G. 2011. *Global Warming Gridlock: Creating More Effective Strategies for Protecting the Planet*. Cambridge: Cambridge University Press.
- Victor, David G., Karl Raustiala, and Eugene Skolnikoff, eds. 1998. *The Implementation and Effectiveness of International Environmental Commitments: Theory and Practice*. Cambridge, MA: MIT Press.
- Wettestad, Jørgen. 2011. "Reducing Long-Range Air-Pollutants in Europe." In *International Environmental Agreements: An Introduction*, ed. Steinar Andresen, Elin Lerum Boasson, and Geir Hønneland, 23–37. Abingdon: Routledge.
- Young, Oran R. 1982. *Resource Regimes: Natural Resources and Social Institutions*. Berkeley, CA: University of California Press.
- Young, Oran R. 1991. "Political Leadership and Regime Formation: On the Development of Institutions in International Society." *International Organizations*, 45(3): 291–308.
- Young, Oran R., ed. 1999. *The Effectiveness of International Environmental Regimes: Causal Connections and Behavioral Mechanisms*. Cambridge, MA: MIT Press.
- Young, Oran R. 2001. "Inferences and Indices: Evaluating the Effectiveness of International Environmental Regimes." *Global Environmental Politics*, 1(1): 99–121. doi:10.1162/152638001570651.
- Young, Oran R. 2003. "Determining Regime Effectiveness: A Commentary on the Oslo–Potsdam Solution." *Global Environmental Politics*, 3(3): 97–104. doi:10.1162/152638003322469295.
- Young, Oran, Leslie King, and Heike Schroder, eds. 2008. *Institutions and Environmental Change: Principal Findings, Application, and Research Frontiers*. Cambridge, MA: MIT Press.