



THE FRIDTJOF NANSEN INSTITUTE

**Case studies of the effectiveness of
international environmental regimes:
Balancing textbook ideals
and feasibility concerns**

Steinar Andresen and Jørgen Wettestad

FNI Report 19/2001



FRIDTJOF NANSENS INSTITUTT
THE FRIDTJOF NANSEN INSTITUTE

Tittel/Title Case studies of the effectiveness of international environmental regimes: Balancing textbook ideals and feasibility concerns	Sider/Pages 21
Publikasjonstype/Publication Type FNI report	Nummer/Number 19/2001
Forfatter(e)/Author(s) Andresen, Steinar and Jørgen Wettestad	ISBN 82-7613-423-8
Program/Programme	ISSN 0801-2431
Prosjekt/Project	
Sammendrag/Abstract This report deals with the principle challenges of case study research on the effectiveness of international environmental regimes related to such issues as case selection, causation, data gathering, regime linkages and qualitative versus quantitative approaches. The main section of the report is a chronological journey through seven projects dealing with various aspects of regime effectiveness, relating experiences to the principal methodological issues and challenges. A central conclusion is that most often there is a trade-off to be made between textbook ideals and feasibility concerns. However, by careful deliberation of theoretical and methodological challenges in the early project phase, combined with an open discussion of the concessions to feasibility, the final trade-off may still produce both interesting and reliable results.	
Stikkord/Key Words regime effectiveness, theory and method	
Bestilling til/Orders to: Fridtjof Nansen Institute, Postboks 326, N-1326 Lysaker, Norway. Tel: (47) 6711 1900 Fax: (47) 6711 1910 Email: sentralbord@fni.no	

Contents

- 1. Introduction 1
- 2. “Textbook ideals”: a brief summary of some principal challenges..... 3
- 3. “Ineffective struggling with effectiveness?” Some main practical challenges..... 5
 - 3.1. Point of departure: ignoring Oran’s caveat...5
 - 3.2. Overview of relevant effectiveness projects5
 - 3.3. “The initial Miles project”: where it all began...6
 - 3.4. Domestic implementation and behavioural change: the IIASA/IEC project and the Hanf/Underdal acid rain project.....8
 - 3.5. “High ambitions”: two projects addressing the complex issue of institutional design. 11
 - 3.6. The Miles Project, phase II (1994-2000): “Confronting Theory with Evidence”. 14
- 4. Conclusion: Balancing textbook ideals and feasibility concerns 17
- References 19

1. Introduction¹

In the paper presented at the conference on ‘Methodological Approaches to the Study of Regime Consequences’ we focused mainly on our experiences acquired over more than a decade as case study workers on the effectiveness of international environmental regimes. Although we have not been among the chief architects conceptually or theoretically in this field, we believe that few others have as much experience in testing out the fruitfulness of various perspectives through case study work. It could be useful, therefore, to sum up in an explicit manner some of the lessons we have learnt on this long journey. Moreover, we believe that an *interaction* between “theorists” and “practitioners” who try out the theories is crucial in advancing the study of regime effectiveness.

In section two of this report we will deal with the principal challenges of case study research in this field, related to issues such as case selection, causation, data-gathering, regime-linkages and qualitative versus quantitative approaches. The third and main section of this report is a chronological journey through seven projects dealing with various aspects of regime effectiveness. As far as possible our experiences are related to the methodological issues and challenges identified in section two. This review places more emphasis on the more or less conscious methodological choices than on the actual findings of the projects discussed. In this account we will also reflect on how some related projects have dealt with these common challenges.

In the final section (four) we will contrast our approach with the “textbook ideals” introduced in section two. In other words, according to the textbook how *should* we in fact have conducted our research? As will be seen, there is a considerable discrepancy between these two approaches. The question is, however, how close can we get to the ideal under real-world circumstances, and what can we gain in extra insight? If we were to start again, would we have conducted research in a totally different manner? What are the trade-offs between the ideals and real-world constraints such as time and costs and finally, what are the implications of deviating from the textbook wisdom?

¹ This report is a revised version of a paper presented at the workshop on ‘Methodological Approaches to the Study on Regime Consequences’, held at the Centre for Advanced Study, Oslo, November 19-20, 1999. Thanks to Rigmor Hjorth and Anne Christine Thestrup for language and editing assistance.

2. “Textbook ideals”: a brief summary of some principal challenges.²

In this section we will briefly address certain principal challenges confronted by every researcher entering the complex terrain of regime effectiveness. The first challenge pertains to the issue of *case selection*. Careful selection can help us maximise what Bernauer and Mitchell (1998) call “internal” and “external validity”. Internal validity involves making “within-case” causal relationships as plausible as possible. External validity has to do with the explicit identification of “the boundaries between the population of cases with which the findings can be validly generalised and beyond which valid generalisations are unlikely” (op. cit. p. 8). Since high internal validity can be seen as a necessary condition for high external validity, and since there is often a need for a trade-off between these ideals, much can be said for letting internal validity take precedence over external validity. How, then, are cases to be selected in order that internal validity be maximised? Some general principles can be outlined: first, one should focus initially on the theory and select empirical cases later. This will contribute to a reduced risk of achieving biased results. Second, look for cases where the values of the independent variable(s) vary, and third, look for more cases than explanatory variables. This makes it easier to hold specific, exogenous (control) variables constant and generally enhances analytical oversight and control. A fourth important recommendation is to try to find particularly “difficult” or “easy” cases in which the control variables either make it very unlikely or very likely that the explanatory variable will produce the theoretically predicted value of the dependent variable.

Thus, although selecting cases to hold the value of certain variables constant increases internal validity, it also automatically limits the range of cases to which one can validly refer (external validity). This paradox can be reduced and external validity increased by using the “difficult”/“easy” case logic mentioned above. Moreover, additional case studies can be conducted in which the control variable has a different value. Finally, by linking own cases and research to past and ongoing research (programs), the logic of literal or theoretical replication can collectively produce results with a wider relevance.

In addition to selecting cases and key variables, a second principal challenge pertains to *operationalisation and data gathering*. A ground rule here is to record and report explicitly and openly the process by which data is generated in order to make data and analyses as replicable as possible. Variables should be defined and operationalised so that the data relate to the theoretical constructs as accurately as possible. However, because appropriate, reliable and observable indicators of complex conceptual variables often prove difficult to find, and this is certainly the case in the regime effectiveness field, it is often necessary to identify various indicators and multiple proxies. In order to give such indicators substance, various types of data should be collected and utilised. This includes reviews of primary and secondary literature, structured or open interviews and surveys, direct or participant observation, or the collection of quantitative data. In practice, due to resource and time constraints, researchers will often rely heavily or entirely on some of these types of data and the strengths and weaknesses of such data “skews” should be explicitly discussed.

² This section draws first and foremost upon King, Keohane and Verba (1994) and Bernauer and Mitchell (1998).

When cases have been selected, hypotheses formulated, and data gathered in the process of analysis, it is a central challenge not only to establish correlation between variables but also to make *sound causal inferences*. This was touched upon in the discussion on case selection, where such selection can serve to strengthen internal validity. But case selection is only part of the picture. Several other aspects and tools need to be mentioned. Central tools in going “beyond correlation” are process-tracing of causal pathways and the careful examination of rival hypotheses. The first of these closely related tools is, by nature, qualitative and not least useful in tracing the paths of institutions’ influence on regime effectiveness. The second tool can be utilised either in a qualitative and intensive manner, or in a quantitative manner (relying on statistical techniques). Used qualitatively, the method of counterfactual reasoning can be helpful. For instance, had a regime commitment not been adopted, would other factors like economic fluctuations and energy switching nevertheless have led to reduced emissions?

This touches on an ongoing discussion among students of international relations relating to the use of *qualitative* versus *quantitative* methods. In their seminal contribution on the issue, King, Keohane and Verba (1994) advocate the many merits of the quantitative approach. Others argue more strongly in favour of the qualitative approach. If conducted properly, the key challenge of trying to demonstrate causation may well be done by the soft-case study approach.³ One illustrative example of the applicability of the case-story approach under real world circumstances is the practice used in a court of law. “In such proceedings judges or juries are asked to make judgements about causation and intent based quite literally on a single case.” (McKeown, 1999:167). To a large extent, the approach used depends upon what you want to know more about and both approaches have shortcomings as well as assets.

A final methodological challenge refers to *the unit of analysis*. Studies of the effectiveness of international regimes will easily zoom in on how individual regimes deal with individual problems. Understanding this is a tall order in itself. However, as environmental problems are often interrelated in intricate ways, and policies adopted in one regime context affect policy-making in other regime contexts, a narrow one-regime – one-problem focus may easily lead us to ignore important questions of interplay, context and linkages. (Young, 1999b, Stokke, 2000)

In the next section we will relate these methodological challenges to the manner in which we have conducted our work on effectiveness of international environmental regimes.

³ See for example McKeown, (1999).

3. “Ineffective struggling with effectiveness?” Some main practical challenges

3.1. Point of departure: ignoring Oran’s caveat...

In 1982 Oran Young wrote: “... there are severe limitations to what we can expect from efforts to evaluate regimes in terms of the outcome they produce... this suggests the importance of giving some consideration to non-consequentialist approaches to the evaluation of regimes.” (Young, 1982:138). In a recent review of the status in this research field, he puts far more emphasis on remaining challenges and uncertainties than on the considerable achievements made so far (Young, 2000). Considering this status after a decade of research by a number of able scholars, maybe he was right? Maybe all of us should have followed his early advice and never embarked upon this avenue. For our part, we are glad that he, as well as the rest of us, did *not* listen to this early word of caution. Our knowledge today is perhaps not as extensive as we expected naïvely a decade ago, but we surely know a lot more than if we had never embarked upon this effort. For example, we are now in a much better position to correct someone unqualified who boasts of the success of a regime. Even if our optimism has been naïve, it has also been productive. If nothing else, it has got us involved in a number of interesting research projects, as will be seen below.

3.2. Overview of relevant effectiveness projects

We have been involved in the following projects that deal, more or less directly, with the effectiveness of international environmental and resource regimes.

1. Miles et al. Phase I 1989-90, “Science, Technology and International Collaboration: Conditions for Effective Global and Regional Action Concerning the Problem of Global Climate Change”; not finished as planned.
2. Related to Miles et al I: FNI effectiveness project (with Underdal) 1990-92 (with some later outputs).
3. Project at the International Institute for Applied Systems Analysis (IIASA) on The Implementation and Effectiveness of International Environmental Commitments (IEC), 1993-96.
4. Project led by Underdal and Hanf on the domestic implementation of acid rain commitments, 1993-95.
5. Project on Science and Politics in International Environmental Regimes, CICERO-FNI collaboration, phase I: 1993-94; phase II: 1997-1999.
6. The Effectiveness of Multilateral Environmental Agreements (MEAs), Nordic project, 1994-95.⁴
7. Project on the Designing of Effective Environmental Regimes, carried out by Wettestad, on and off 1993-98.
8. Miles et al. phase II, 1994-2000.⁵

⁴ The final publication from the project is Stendahl-Rechardt et.al. (eds.), (1996). We covered CLRTAP (Wettestad) and the North Sea (Andresen). We learned a lot about the differences between international lawyers and political scientists concerning how to deal with effectiveness, but since we did not gain much additional insight compared to other projects we have participated in, we will not discuss this project here.

⁵ FNI has also conducted or been involved in a number of other effectiveness projects that will not be discussed here. E.g. Skjærseth (2000), Rosendal (2000), Stokke and Vidas (1996) and Young et.al. (1999a).

Almost all our work within these projects has been case-study oriented. Arild Underdal has provided most of the intellectual capital to the majority of these projects. While he discusses the overall perspective in this book, a more practical ‘street-level’ story may bring some useful complementary insights.

3.3. “The initial Miles project”: where it all began...⁶

It all started with a project initiated by Professor Ed Miles called “Science, Technology and International Collaboration: Conditions for Effective Global and Regional Action concerning the Problem of Global Climate Change”.⁷ Two main project phases were envisaged, with the first one summing up main lessons from the field, and the second and main phase applying these lessons to the then fresh and exciting issue of how to address global warming. Based on these lessons, how should the climate regime be designed?⁸ In terms of the key concept of effectiveness, although the project title talked about “effective action”, it seems right to say that we – like virtually all others at the time – started out with a perspective geared largely towards *outputs* produced by regimes. This probably explains why “level of collaboration” figured as a central dependent variable in the early project talks. Much time within this project was spent on achieving a shared understanding of the basic model – i.e. main variables, relationships between variables, the formulation of hypotheses, and how the variables and hypotheses were based on various strands of International Relations theory.⁹ The only methodological issue, in a strict sense, that was discussed in some depth during the initial project meetings referred to *case selection*. The need to cover as many and varied cases as possible was stressed by the initiator of the project. The main criteria for assigning research work were existence and availability of competence: which regimes had the various collaborators already conducted work on or had fairly easy access to relevant data on. Feasibility and pragmatism were the keywords.

In line with textbook ideals, we started out with a main emphasis on theory. However, with the benefit of hindsight, we did not pay enough attention to the link between theory and case selection, apart from aiming at a large number of (varied) regimes. Theoretically, as well as empirically, the project was very ambitious, as is often the case when real-world constraints do not receive sufficient attention. As it turned out, the original Miles et al. project did not obtain US financial support.¹⁰ A more modest and simplified version of the *first* part of the initial project is now near completion (see Miles et al. II) – more than a decade after it started...

For their part, the FNI team obtained a modest two-year grant to start on a simplified version of the eclectic and comprehensive analytical approach presented in the common project note.

⁶ See also Andresen and Wettestad (1995).

⁷ Andresen spent his sabbatical at the University of Washington at the time, thereby facilitating the link to the FNI.

⁸ An early attempt at applying this perspective was reported in Andresen and Wettestad (1992).

⁹ As this was Professor Miles’ baby in the first place, he spent much time in the cupboard at the first meetings; “drawing and explaining” while the FNI delegation listened and made notes.

¹⁰ For a fascinating account of the history of this (and other) Miles’ projects, see Miles (1997). Suffice it here to explain that one reason why financing was not obtained was the link to these ‘environmentally radical’ Norwegians. Recall that the official US climate policy at the time was not very progressive. See Agrawala and Andresen (1999).

The basic conceptual work was developed by Arild Underdal in several draft notes and the main ideas were circulated as a working paper in 1990. The dependent variable (“effectiveness”) had here been developed further than in the initial project note and the independent variables had been simplified through the two key concepts of “problem structure” and “problem solving capacity”.¹¹ This seminal model and concepts have no doubt continued to influence much FNI work on this subject up to the present. Moreover, we think it is also safe to conclude that these basic early ideas have had significant impact over time on the thinking of large parts of the “effectiveness community”. Again, the textbook was followed by the elaboration of theory as a first step, and this time a seemingly simpler and more unified approach paved the way for the empirical research.¹² However, when we started selecting cases in 1990, we sinned against the textbook ideal again, since our case selection primarily grew out of the pragmatic thinking of the Seattle-Oslo project context. It was *not* extensive discussions over internal and external validity that formed the basis for selecting the five cases to be studied: The whaling regime, the acid rain regime, the Paris Convention on land-based marine pollution (PARCON), the Oslo Dumping Convention (OSCON), and the international regime to conserve the marine resources in the Southern Ocean (CCAMLR). The first three were selected because we were very familiar with these while the two remaining regimes were new to us. In general, not much systematic research had been undertaken on these regimes at the time. Moreover, we saw some interesting differences between them. For instance, two of them dealt with marine living resources, three of them with the environment; two of them were global and three were regional; some were rather inaccessible and “closed”, while others were more “open”; and we also expected disparities in terms of effectiveness.

The issues of data collection and causation are closely linked in this case since real-world working conditions and perhaps intellectual limitations did not allow much in-depth case study research.¹³ Somewhat paradoxically, a master’s dissertation on the (North Sea) Oslo Commission was the only case study that could fully take on the formidable empirical and analytical challenges presented in the analytical structure. In terms of main sources and data, the authors of this chapter largely used second-hand sources; primarily books and articles. Although some first-hand sources from regime meetings etc. were consulted, this vast array of information was not systematically utilised. These sources were complemented by a “semi-structured” questionnaire sent out to “expert panels” consisting of regime parties and a select group of more independent observers (including NGOs and scientists).¹⁴ In addition, a few interviews were conducted. Maybe the single most important shortcoming was that most cases in this early phase were not detailed enough to allow us to address the crucial question of causation in a fairly meaningful manner.

¹¹ Since we assume these concepts are explained by the founding father in this book, we see no reason for a more systematic elaboration here.

¹² The FNI team was strengthened by Jon Birger Skjærseth, a graduate student at the time. He has been on our team off and on throughout the entire period and concluded his Ph.D thesis on the North Sea co-operation in 1999.

¹³ Given FNI’s very applied nature (at the time) and soft financing structure, we were never able to concentrate our full attention on this project. Still, and maybe fortunately, strict financial project accounting had not yet been introduced at the FNI, allowing us to spend more time on this project than we actually had funding for.

¹⁴ In the case of CLRTAP we received eight replies from the group of parties and four from observers. The response was good as regards the IWC, but was not much use due to the polarisation of the whaling issue.

Finally, a few words on some of the problems we encountered when trying to apply the theoretical approach. Although the approach was elegant as well as useful, all concepts were not easy to apply. This was especially true for one of the effectiveness indicators, “distance to collective optimum”. As has been pointed out quite recently, this criterion is not easy to apply, even under ideal circumstances (Young, 2000), and with our limited data, application was somewhat random and ad hoc. Another unperceived aspect of the theoretical approach was that it devoted little explicit attention to the role of domestic institutions and policy.¹⁵ This was possibly due to the fact that Underdal’s paper was geared towards understanding the negotiation of effective solutions, naturally downplaying the role of the domestic scene (Underdal, 1990). But other factors also contributed to our “domestic neglect”. Important implementation processes had just started in several of the regimes we studied. As some of these regimes were less than a decade old, the ink had only just dried on important protocols and declarations. Thus, some of the cases were premature in terms of tracing behavioural impacts – which of course should have been given more consideration when cases were selected. With the benefit of hindsight, with this rather slim empirical background and the short time elapsed, it did not make much sense to bring in the even more complicated impact issue, i.e. whether the regime in question had an impact on the problem at hand. But this is exactly what we did. In fact, since we more or less skipped the domestic scene, we moved directly from output to impact.

With regard to explaining effectiveness, we assumed that problem structure was the most important perspective. This perspective contained many sub-dimensions that required considerable analytical attention. The issues of problem-solving capacity and institutional design, more analytically interesting as such factors may be more easily changed and manipulated, were overall given a short shrift in this first phase. In our “summing up so far” article from 1995 (Andresen and Wettestad, 1995), we emphasised the need for more attention, not least to the domestic processes but also to the questions of institutional design and problem-solving capacity. At least four projects were underway at this point which addressed these concerns and in which we were involved. Let us first briefly sum up two “domestic” projects before reviewing two “institutional” projects.

3.4. Domestic implementation and behavioural change: the IIASA/IEC project and the Hanf/Underdal acid rain project.

It has been maintained that the study of domestic implementation of international environmental commitments is a side-track to the study of effectiveness of regimes (Young, 1998). In our opinion this depends on how one defines domestic implementation. If it is perceived as the formal process of transforming international rules into national legislation, we agree. If, however, target groups define it as regime-related behavioural change this is at the very heart of effectiveness studies.¹⁶ In fact, defined this way, we will argue that in order to conduct process tracing, find causal pathways, and undertake counterfactual analysis, the detailed study of domestic implementation processes offers an excellent opportunity. Since we had more or less neglected these processes in our first-generation effectiveness studies, the chance to participate in two “implementation” projects, the IIASA project and the Hanf/-

¹⁵ It was Skjærseth’s (1991) more detailed study that first brought this problem to the fore.

¹⁶ In our work we have focused on target groups in a narrow sense. Others have defined it more broadly, see Stokke and Vidas (1996).

Underdal project, offered a promising opportunity to delve deeper into our understanding of effectiveness from this perspective.

Turning first to the project led by Ken Hanf and Arild Underdal, we have less to add here than for some other projects we have participated in, since we were not much involved with designing the theoretical framework or with selecting the cases. From our perspective, the project was sound, fairly straightforward and simple from a methodological point of view. There was one key regime – the acid rain regime: what effect did this regime have on domestic implementation in ten selected member countries? The acid rain regime stood forth as a natural choice, considering it was comparatively mature. The first protocol was adopted back in 1985; quite a lot seemed to have been accomplished by the regime and new protocols were being added. This rather narrow approach in terms of issue area and regime implied that some of the “big” questions we had dealt with within the multi-regimes projects, such as the comparison between regimes and the more complicated issue of case selection, were not touched upon. This is *not* to say that the approach adopted within this project was smooth and easy to deal with in terms of analytical perspectives and method. One key issue within the project was whether a quantitative or a qualitative approach should be adopted.¹⁷ In the end, the final product was essentially of a qualitative nature.¹⁸

We think that there is much to say for this rather simple approach. Since there was *one single* regime common to as many as ten countries, and the case studies were mostly carried out by scholars from the country in question, fairly solid cross-country comparisons were possible. More important from our perspective was the possibility to study the issue of causation in more detail, i.e. the causal significance of the regime for the implementation process in each country. This approach contributed more than most effectiveness projects we are familiar with in explaining the effectiveness of a single regime in terms of behavioural change. By using rival hypotheses, we were able to demonstrate in the case of Sweden and Norway (for which we were responsible), that the regime had not been the main driving force in the initial reductions in SO₂ emissions, but sooner unrelated factors.¹⁹ In the case of Norway, it was concluded that “... a closer look at the measures reveals that only a few of them can be said to have involved any cost to Norway, and they were not introduced as result of international agreements, nor motivated by acidification”.²⁰ This finding would have been hard to discover without detailed process tracing based on high-quality case study work.²¹

The main weakness of this project was not related to what was achieved, but to the missed opportunity to conduct a closer investigation of inter-linkages between related institutional contexts, i.e. the European Union (EU) and the Convention on Long-range Transboundary Air Pollution (CLRTAP). For various reasons this was not pursued as envisioned in the application. Some of the case studies indicated that for some (laggard) countries, EU

¹⁷ On a continuum from soft to hard, Dr. Sprinz was at one end and Dr. Boehmer-Christiansen on the other while Professor Underdal was the chief mediator as he had a foot in both camps.

¹⁸ This applied to the report sent to the Commission, as well as the later book version (Hanf and Underdal, 2000).

¹⁹ Because Norway had made such progress in emissions reductions, they had a moral upper hand and were quite successful in using CLRTAP strategically to push more reluctant parties in the initial phase. (Laugen, 1995)

²⁰ Hanf et.al (1996:75).

²¹ We can say this since we were not the ones who actually wrote the reports. This was done by two very talented young scholars from Norway and Sweden.

Directives were more important than CLRTAP commitments. Elaborating more systematically on the links between the two regimes could have been one of the first main contributions in this field. Subsequently the links between the EU and CLRTAP on this issue have only grown in importance (Wettestad, forthcoming 2002).

The project carried out at the Institute for Applied Systems Analysis (IIASA) was much more than case-study work. Still, it seems fair to say that this was the major part of the project and it constitutes the main core of the IIASA effectiveness project book.²² (Victor et al., 1998). In terms of what we should do and how we should do it, this was a much more open-ended and process-oriented project than the acid rain project. This certainly made the initial, rather long, process exciting since we all had some influence on where we were moving.²³ The project application made it clear that a main module should deal with domestic implementation, but not much was said about particular regimes, countries and more specific approaches.

Turning first to the question of case-selection, this was finally done in the same way as in all international research projects that we have participated in: rather pragmatically – based on expertise and preferences of the team members.²⁴ Since IIASA is a more political organisation than most other research institutions we are familiar with, there was an East-West dimension to the project, affecting approaches and case selection. Initially a much more “ideal” approach was conceived, but as usual time went faster than expected and money was in shorter supply than envisaged at the outset. Thus, we all ended up writing accounts about the cases that we were familiar with; the acid rain regime, the whaling regime, the London dumping regime and the North Sea regime. The Russian case study participants chose a similar approach.²⁵

Although this is not according to the textbook, there may also be certain advantages in applying such a pragmatic approach, since it increases the chances of achieving rather good case-stories. By and large we think this was true of the IIASA project. Thus, causation, process-tracing, as well as rival hypotheses were discussed fairly systematically in most cases. Moreover, there is no doubt that the Baltic regime²⁶, the North Sea regime and the acid rain regime are appropriate for studies of domestic implementation. It is more doubtful whether the same goes for the whaling regime.²⁷ It may also be argued that the three regional regimes mentioned are a bit too similar along many dimensions. All of them are regional regimes, with fairly limited participation and involving only developed countries. Two of them include

²² We are mainly addressing the final main part of the book, dealing with domestic implementation.

²³ Some 15 researchers from a number of countries were involved in the project.

²⁴ We are not familiar with all major research projects carried out over the last decade. The selection process may have been much more ‘ideal’ in the ‘Tubingen-project’ and the ‘Jacobsen – Brown Weiss’ project, as we know little about the internal life of these. Although we have not been a part of the Young et.al. effectiveness project (Young, 1999), based on the expertise and interests of the members and the cases covered, they seem to have used a pragmatic approach as well.

²⁵ The same procedure was used to some extent in the other main part of the project dealing with international review mechanisms, (IRM) but due to the tremendous work capacity of Dr. Victor, this team may have been able to use a somewhat more ideal approach.

²⁶ See Roginko (1998) and Hjort (1998).

²⁷ Since Andresen wrote this piece, we are in a position to say this without hurting anybody’s feelings. The reason why the whaling regime is not so suitable for such studies is that there was not much domestic implementation in the traditional sense. As the situation was at the time, the cases were selected too late in the process to have done this differently.

both Eastern and Western European countries. Some more variation would have added more analytical mileage. Considering the resources and the personnel involved in this project, it should in principle have been possible to initially choose a more ideal strategy from a methodological perspective.

As for the analytical approach, we were searching for a niche, to make it easier to produce a simple “take home” message. After long discussions in the hot summer of 1994, we ended up with *access and participation*. According to one of the reviewers of the IIASA book, this was not a happy choice, and to some extent we agree. With the benefit of hindsight, it might have been better to basically compare the implementation processes of East and West in a systematic fashion – considering the composition of the teams and the regimes chosen. Still, looking at it more from a process perspective, it did provide us with a more specific focus. It also had obvious connections to very seminal political science debates, touching upon some core democracy issues and discussions. However, within the complex world of “regimes, states and societies”²⁸, the approach sort of melted in our hands – although subsequent summers were not quite that hot. By some, we were (rightfully) accused of providing “waffling” messages – and we (rightfully) responded that the access and participation issue was a waffling one. A simpler approach would probably have functioned better for the majority of case study authors who could not link the IIASA work to deep-diving dissertation work.²⁹ But that doesn’t mean that our work didn’t produce interesting results. For instance, although the IWC was no ideal case, the work on this regime questioned several generally held assumptions in this field; both that maximum transparency and NGO involvement is always good, and that NGOs and states are always on the opposite side of the floor. Finally, work within the IIASA project provided helpful inputs to other projects that we were involved in at the time, such as the project on the science-politics interface, Wettestad’s institutional design project as well as Skjærseth’s Ph.D study on the North Sea.

3.5. “High ambitions”: two projects addressing the complex issue of institutional design.

Wettestad and Andresen (1991) had already concluded that there was a need for closer study of problem-solving capacity and especially the intriguing question of institutional design. Based on a couple of preliminary papers/articles written jointly by Andresen/Wettestad,³⁰ Wettestad started on a more comprehensive “institutional” project during his 1993 sabbatical. In order to benefit from and provide further inputs to the Miles/Underdal effectiveness project, both the dependent variable (“effectiveness”) and the “problem structure” explanatory perspective were quite similar to the “Miles/Underdal” model. The main differences lay in 1) the institutional specification; and 2) the treatment of the problem structure perspective as a “rival hypothesis” control perspective. With regard to the institutional specification, six main institutional variables were identified and discussed in the 1995 preliminary project note (Wettestad 1995): access procedures and participation issues; decision-making rules; the role of the secretariat; the structuring of the agenda; the organisation of the science-politics

²⁸ These were the three key words in the so-called ‘Think Piece’ produced by the core FNI team. (Andresen et al. 1995)

²⁹ Skjærseth was in such a position, and he proved that it might not have been the niche that was wrong but the ability to ‘deliver’ on the part of the rest of the team.

³⁰ Andresen and Wettestad, (1992), (1993).

interface; and verification and compliance mechanisms. Case selection was again done pragmatically, based on what could be utilised from own research and from other contributions from the FNI. Eventually, the project ended up by focusing on three main cases (OSCON/PARCON; CLRTAP; ozone regime) and a fourth, rough analysis of the organisation of the climate regime so far (Wettestad 1999). Although these cases offered a certain variation both in terms of effectiveness and the focused institutional variables, the selection of other cases could undoubtedly have increased this variation.

On the institutional side, the main new source of information in addition to earlier efforts was a series of institutionally focused interviews with central Norwegian participants in the regime processes under scrutiny. Although ideally far more interviews – and in several key countries – should of course have been carried out, given the complex character of the institutional issues, the author was quite satisfied with this “fewer, but deeper” approach as regards the interviews. In terms of overall methodology, the high ambitions of the project were reflected in an effort to carry out a systematic institutional comparative effort. Although comparing three/four environmental regimes may not sound like a tall order to some, for one researcher alone it turned out to be a very tall order indeed.³¹ One basic problem, of course, is that in order to carry out meaningful comparisons, you first need to have a pretty good understanding of the components you are seeking to compare. Take, for example, the decision-making/-ratification requirements within the ozone regime: It is hard enough to grasp all the nuances in texts and differences between the procedures related to the Convention, Protocol and amendments. Add then, first, the task of verifying the practical, causal impact of these nuances and differences in the vast number of processes and for the large number of actors. A safe bet for political scientists is “not very much practical impact”, but safe bets may sometimes turn out to be terribly wrong. Then, add the need to understand the operation of the decision-making procedures in relation to the rest of the complex ozone regime machinery. Hence, the importance of *context* and *conditioning factors* became the main, general lesson that stood forth from the author’s efforts to juggle all these processes, actors and procedures. There is no “effective regime design” as such. Factors like problem types (e.g. conflicts over interests versus conflicts over values), phase in the development of the regime (e.g. the early confidence-building phase versus the more mature, “confident” phase) and process types (e.g. processes of preparation versus negotiation versus implementation) need to be taken into consideration.

The other institutionally-focused project started in 1993 as a collaboration project between CICERO and FNI on the organisation of the *science-politics interface in international environmental regimes*. Given the many complexities related “only” to this institutional issue, this alone indicates the folly in taking on the vast agenda indicated above. As to analytical approach, we singled out one small piece of the explanatory perspective used in the broader effectiveness projects. Regarding the dependent variable we used a three-level indicator related to acceptance and adoption of scientific input. Based on previous research, we did not expect scientific input to be among the main premises determining policy. A wide range of other factors may be more important. We therefore decided to check the significance of some other key variables: i.e. state of knowledge, political malignancy and public saliency. Again,

³¹ However, projects with *several* researchers addressing the complex institutional issues automatically run into problems of inter-subjectivity with regard to interpretation of concepts, and measurements and assessments

we think our analytical approach was quite solid, based on a template mainly produced by Underdal, but the team effort was central for this result as well as the final outcome.³²

But the link between the theoretical point of departure and case selection was basically pragmatic. Due to previous work and the case study work within the Miles/Underdal context, it was again tempting to “bet on the old horses” for the FNI-team, which meant PARCON/North Sea Conferences, CLRTAP and IWC/whaling. Tora Skodvin from CICERO brought some fresh ideas to the group, and added the climate change regime as a new case in addition to the ozone regime. But selection was equally pragmatic on her part.³³ Nevertheless, again we expected a significant variance between the regimes on a number of dimensions, including the importance of how the scientific input is organised. As to data gathering, we could to some extent build on previous science-policy research carried out at the FNI (e.g. Andresen/Østreng, eds., 1989; Wettestad, 1989, Andresen, 1989). That is, we were already familiar with the broad lines of the science policy interplay within the regimes. Still, some time had elapsed and the science-politics relationship was by no means static. Moreover, as this project was geared more specifically towards the *institutional design* issue, it was more demanding than our previous research on the general relationship between science and politics. This implied that we had to collect new data, not least through a series of interviews with scientists and policy makers within the different regimes.

During the final phase of the project (1997-99), we think it’s fair to say that we acquired more and better data, although there was some variance between the cases in this regard.³⁴ The most important improvement from a methodological point of view was the combination of fairly detailed case studies and a more systematic and quantitative comparative effort. Others must judge whether this was successful, but we think it made good sense to contribute to bridging differences between various academic traditions. We also think that such an approach gives added insight and forces us to be more transparent as well as systematic.

As to main findings, we concluded that, overall, the state of knowledge was more important than the design of this process. Still, the institutionalisation of the science-policy interface made a difference under certain circumstances. As expected, institutional design within the regimes was geared towards maintaining a certain balance between scientific integrity and involvement/feasibility. However, the balance was more tilted towards involvement than we expected. But the fact that advice was not always strictly independent caused fewer problems than we had predicted. This was perhaps not much of a general take-away message, but such a message is not always easy to find when one attempts to delve deeply into a complex issue. A simple, blunter, message would be that knowledge often seems to mean little and how the knowledge-making process is designed means even less to the effectiveness of international regimes, but this is also not much of a “take-home” message.

³² The final product is now published, Andresen/Skodvin/Underdal/Wettestad, (2000).

³³ Skodvin worked on her Ph.D thesis on IPCC/the climate regime (Skodvin, 1999).

³⁴ Personal attendance at some of the relevant negotiation meetings was an additional important source of data.

3.6. The Miles Project, phase II (1994-2000): “Confronting Theory with Evidence”.

Several factors have made it possible for us to improve our “Miles” case studies during the second and final project phase. Since we had worked on most of the cases for a long time, we now had a better knowledge of our cases.³⁵ The work of others, both conceptually and empirically, has also contributed to improving our understanding. We were therefore in a much better position to estimate the real causal effects of the regimes, compared to other factors. The process by which data-gathering took place and causation was discussed, had improved considerably. From a methodological and analytical perspective this is clearly the most ambitious effectiveness project we have been involved in, both in scope and depth. Nevertheless, the pragmatic approach has essentially been used during case selection, building on the first phase of the project. But the number of cases is considerably higher than in other projects we are familiar with. The number of regimes is 14 (FNI 7 regimes and Miles et al. 7 regimes), but the number of cases is “thirty-something” as the cases are split into regime phases and components.³⁶ Still, even in comparison to a not-so-large universe, the number of cases is far too small and not meant to be representative. Hence, no substantial conclusions in a statistical sense can be drawn. But compared to other effectiveness projects, there is a rather unique combination of significant quantitative analytical element and many in-depth case studies. Since the standardisation of the cases has improved considerably, we think this project will make a significant contribution to bringing research on the issue one step forward.³⁷ As case study workers we find the “hard-soft” approach appealing, although we have yet to tap the full potential of the aggregate findings from the soft approach. Still, compared to our feeble start, we are now in a better position to shed light on why some regimes fail and others succeed.

After these somewhat self-congratulatory remarks, it is time to make the reader aware of some of the problems that still remain. Young (2000) is right when he remarks that the subjective judgement of the individual analyst when explaining and evaluating the effectiveness of various regimes is crucial. This is a weak point that is hard to get around. The problem is compounded when the project has stretched out so long and the number of analysts has been fairly large. We have tried to control for this by intensive group coding, but are far from certain that we have been successful. Cultural differences, personal opinions, preferences, styles, experiences and biases are bound to have an effect. But at least we are transparent and present all our scores in a separate appendix.

A final point on our (semi-) statistical approach: in principle there should be nothing mysterious about applying numbers to something that has already been described in prose. Nevertheless, from the perspective of those having to decide whether the score on some complex dimension should be “two” or “three” on a five-point scale, this can be a very painful process. Maybe it’s because we belong to a more qualitative tradition that we find it much more comfortable to describe and analyse in prose. In addition, the increased transparency

³⁵ We guess a cynic’s comment would be: “that’s about time!” And that may be true, but the more we work on these regimes the more we understand how complex they are.

³⁶ Not all the reviewers in the last round (i.e. spring 2000) were happy with our many phases and regime components, but we consider this approach to be essential to analyse the evolving nature of regimes.

³⁷ Some of the reviewers did not find the standardisation good enough in the first round. We agreed and have improved it, and overall we think the standardisation is better than in most comparative projects.

related to quantification makes us all the more vulnerable, since we know there are many out there who disagree with us – often with good reason. You will now all be able to judge what score you will assign to our main product (Miles et al., 2001).

4. Conclusion: Balancing textbook ideals and feasibility concerns

Let us first briefly sum up of our work related to methodological approach. Generally, we have been true to the textbook ideal by starting out with a more general analytical approach, but have not been able to follow this up in relation to case-selection. Our most persistent sin has been in the selection of cases. In none of our projects have discussions of external and internal validity formed the basis for case selection. Although our arsenal of cases has expanded over the years, all cases have generally been chosen pragmatically. As for the issues of causation and data collection there have been considerable improvements over time. This, however, has partly been the result of a learning process, as we have become more familiar with some key cases over time. In our opinion neither the qualitative nor the quantitative approach is superior to the other. There is no “correct” method in this regard; they clearly supplement each other – and we need both. At heart we are qualitative caseworkers, but this may be the result of training, lacking skills and tradition rather than of principles. The attempt at bridging the two approaches introduced by professor Underdal has been somewhat painful but also stimulating. Finally, to the question of analysis units and regime linkages and interplay. This is a rather recent fad, but one aspect that we would have been able to explore better at a much earlier stage, if we had been true to our initial focus on complex problems rather than single regimes. However, in a complex reality, we found that regimes were much easier to pin down compared to changing and comprehensive “problem structures”. In the most recent project, however, this broader perspective has been included to some extent.

Turning then to the seminal question of “what would we have done differently had we known then what we know now?”, posed to us by the editors of this book in a review of our original paper, at least two points stand out. First, we should have set aside some (more) time to study and discuss methodology – not a common subject at the FNI a decade ago. Even though time and resources are always limiting factors, this could have reduced some of the initial shortcomings. Second, rather than starting out with five cases, we should probably have started out with one or two. Considering our ambitious analytical approach, with the benefit of hindsight, we really stood no chance of answering the questions we set out to answer for so many cases. The best approach would probably have been to have carefully singled out only one case as a “pilot case” to test our perspective. Then we would have been in a much more favourable position regarding process tracing, domestic implementation, data collection etc. Also, we would have been able to test the fruitfulness of the analytical perspective much more systematically. To illustrate this point, it can be noted that two of the doctoral theses produced at the FNI within the ‘effectiveness and implementation’ tradition in the 1990s both used around four years to study these questions within *one* regime. No wonder then that especially our initial effectiveness research had some of the weaknesses we have shared with you in the previous section.

Secondly, is there a trade-off between textbook ideals and feasibility concerns? In research institutions not part of the university system we think the answer is generally “yes”. But we hope that this chapter has shown that after careful deliberation of theoretical and methodological challenges in the early project phase, combined with an open discussion of the concessions to feasibility, the final trade-off may produce both interesting and reliable results. Although we have been quite humble about the limitations of our findings due to the lack of

methodological considerations, maybe authors of textbooks in methodology should also consider real-world constraints to a larger extent. If not, they may tend to be rather lofty publications without much practical significance for empirical research.

Finally, to the final question raised by the reviewers of the original paper, the implications of deviating from the textbook wisdom. In our case this primarily relates to case selection. We have already touched upon one alternative route: to dive deeper down into one case and do that more thoroughly before expanding to consider new cases. An alternative route would be to start out by mapping the "universe" of international environmental – and resource regimes. Even though this is only a small subset of all international regimes, there are considerable methodological challenges in delimiting them from other regimes as the borderlines between them are often not that clear-cut.³⁸ Leaving this challenge aside, it is quite clear that we now know a lot about a rather small number of regimes. Apart from the ones we have discussed here, the climate regime, the biodiversity regime and a few others are the main candidates.³⁹ Thus, we know little or nothing about the large majority of international environmental and resource regimes. The reason for this selection bias is of course that the regimes which are studied most are considered to be the most important both from a political and analytical point of view. Still, this is troublesome from a methodological perspective as they may not be representative for all cases. The most ambitious effort to correct this is the large-scale so-called "IIASA-database-project".⁴⁰ It has proven very time-consuming and not easy to finance, but the effort should be commended, even though we don't yet know the final outcome.

³⁸ For an overview of the number and growth of environmental regimes compared to other regimes, see Agrawala, 1999 (Ph.D thesis). For a discussion of the difficult borderline between environmental - and other types of regimes, see Andresen, 2000.

³⁹ It should be noted, however, that some of the reviewers of the Miles II project have complimented the fact that Professor Miles in particular has brought in some new cases. See Miles et al., 2001.

⁴⁰ For some information on this project, see Breitmeier et al, 1996.

References

- Agrawala, S. (1999) Science Advisory Mechanisms in Multilateral Decisionmaking: Three Models From the Global Climate Change Regime, ph.d. thesis, Woodrow Wilson School, Princeton University.
- Agrawala, S and Andresen, S. (1999) Indispensability and indefensibility? The United States in the Climate Treaty Negotiations, *Global Governance*, 5, 1999, 457-482.
- Andresen, S. (1989), Science and politics in the International management of whales, *Marine Policy*, Volume 13, Number 2, 99-118.
- Andresen, S, 2000, Global Environmental Governance: UN Fragmentation and Co-ordination, *Yearbook of International Co-operation on Environment and Development*, 2000/2001 (forthcoming), Earthscan, London.
- Andresen, S. and Wettestad, J. (1992). International resource cooperation and the greenhouse problem, *Global Environmental Change*, vol. 2, no. 4, December, 277–291.
- (1993). The effectiveness of international resource co-operation: some preliminary notes on institutional design, *International Challenges*, vol. 13, no. 2, 61–75.
 - (1995). International problem-solving effectiveness: the Oslo Project story so far, *International Environmental Affairs*, vol. 7, no. 2, Spring, 127–150.
- Andresen, S. and Østreng, W. (eds.) (1989). *International Resource Management*, London, Belhaven Press.
- Andresen, S.; Skjærseth, J.B. and Wettestad, J. (1995). Regime, the state and society: analysing the implementation of international environmental commitments, Laxenburg, IIASA Working Paper WP-95-43, June.
- Andresen, S., Skodvin, T., Underdal, A. and Wettestad, J. (2000). *Science and Politics in International Environmental Regimes: Between Integrity and Involvement*, Manchester University Press, Manchester.
- Breitmeier, H. et al. (1996). The international regimes database as a tool for the study of international cooperation, Laxenburg, IIASA Working Paper WP-96-160.
- Hanf, K, et.al., (1996), Final report to the European Commission, EC Contract EVSV-CT92-0185, The Domestic Basis of International Environmental Agreements: Modelling National/International Linkages
- Hanf, K. and Underdal, A. (eds.) (2000). *International Environmental Agreements and Domestic Politics: The Case of Acid Rain*, Aldershot: Ashgate
- Hjort, R. (1998) Implementation of Baltic Sea Pollution Commitments in Poland: A review of the Literature, in Victor, D.G.; Raustiala, K. and Skolnikoff, E.B. (eds), *The Implementation and Effectiveness of International Environmental Commitments*, Cambridge MA, MIT Press, 639-659.
- Laugen, T. (1995). Compliance with international environmental agreements – Norway and the Acid Rain Convention, R:003-1995, The Fridtjof Nansen Institute.

Mitchell, R. and Bernauer, T. (1998). Empirical Research on International Environmental Policy: Designing Qualitative Case Studies, *Journal of Environment and Development*, vol.7, no.1, 1998, 4-31.

King, G.; Keohane, R. and Verba, S. (1994). *Designing Social Inquiry: Scientific Inference in Qualitative Research*, Princeton NJ, Princeton University Press.

McKeown, (1999), Case Study and the Statistical Worldview: Review of King, Keohane and Verba's Designing Social Inquiry: Scientific Inference in Qualitative Research, *International Organization*, 53, 1 Winter, 161-190.

Miles, E.M; Underdal, A.; Andresen, S.; Wettestad, J.; Skjærseth, J.B. and Carlin, E.M. (2001). *Environmental Regime Effectiveness – Confronting Theory with Evidence*, Cambridge MA, MIT Press.

Roginko, A. (1998), Domestic Implementation of Baltic Sea Pollution Commitments in Russia and the Baltic States, in Victor, D.G.; Raustiala, K. and Skolnikoff, E.B. (eds.), *The Implementation and Effectiveness of International Environmental Commitments*, Cambridge MA, MIT Press.

Rosendal, G.K. (2000), *The Convention on Biological Diversity and Developing Countries*, Kluwer Academic Publishers, Dordrecht/Boston/London.

Stendahl-Rechardt, K. et.al. (1996). *The Effectiveness of Multilateral Environmental Agreements*, Nordic Council of Ministers Report 18:1996.

Skjærseth, J.B. (1991). (In Norwegian) Effektivitet, problem-typer og løsningskapasitet: en studie av Oslo-samarbeidets takling av dumping i Nordsjøen og Nordøstatlanteren, R:009–1991, The Fridtjof Nansen Institute.

- (2000). North Sea cooperation Linking international and domestic pollution control Manchester University Press, Manchester and New York.

Skodvin, T. (1999), Structure and Agent in the Scientific Diplomacy of Climate Change, Dr.polit. thesis, Department of Political Science, University of Oslo.

Stokke, O.S. and Vidas, D. (1996), *Governing the Antarctic: The effectiveness and legitimacy of the Antarctic Treaty System*, Cambridge, Cambridge University Press.

Stokke, O.S. (2000), Managing straddling stocks: the interplay of global and regional regimes, *Ocean & Coastal Management* 43 (2000) 205-234.

Underdal, A. (1990). Negotiating effective solutions: the art and science of political engineering, unpublished paper, the University of Oslo.

Victor, D.G.; Raustiala, K. and Skolnikoff, E.B. (1998) (eds), *The Implementation and Effectiveness of International Environmental Commitments*, Cambridge MA, MIT Press.

Wettestad, J. (1989). *Uncertain Science and Matching Policies: Science, Politics and the Organization of North Sea Cooperation*, R:003–1989, The Fridtjof Nansen Institute.

- (1995). *Nuts and Bolts for Environmental Negotiators? Institutional Design and the Effectiveness of International Environmental Regimes*, FNI Note N:001–1995, The Fridtjof Nansen Institute.

- (1999) *Designing Effective Environmental Regimes – The Key Conditions*, Cheltenham: Edward Elgar.
- (forthcoming 2002). *Clearing the Air – European Advances in Tackling Acid Rain and Atmospheric Pollution*, Aldershot, Ashgate.

Wettestad, J and Andresen, S. (1991) The effectiveness of international resource co-operation, Some preliminary findings, R:007, The Fridtjof Nansen Institute.

Young, O.R. (1982). *Resource Regimes Natural Resources and Social Institutions*, University of California Press, Berkeley, Los Angeles, London.

- (1998). The effectiveness of international environmental regimes. what we know and what we need to know, paper produced for the annual meeting of the American Association for the Advancement of Science, Philadelphia, 12–17 February.
- (1999a). Oran R. Young et al, *The Effectiveness of International Environmental Regimes*, Cambridge, Mass.: MIT Press.
- (1999b). Science Plan for the project on the Institutional Dimensions of Global Environmental Change, Bonn: *The International Human Dimensions Programme of Global Environmental Change – IHDP*.
- (2000). Epilogue: Indices and Causal Inferences: Exploring the Effectiveness of International Regimes, in J. Wettestad (ed.). *The Effectiveness of Global and Regional Environmental Agreements: Proceedings from the 1999 Oslo Concerted Action Workshop*, FNI Report 1/2000, The Fridtjof Nansen Institute, 49-62.