

**Developing a Multi-Disciplinary Evaluation of an Environmental Policy Innovation:
Impacts of Forest Certification within the United States**

Benjamin Cashore
Assistant Professor
Chair, Program on Forest Certification
Yale School of Forestry and Environmental Studies
203 432-3009
benjamin.cashore@yale.edu

Emily Noah
Research Associate
Program on Forest Certification
Yale School of Forestry and Environmental Studies
emily.noah@yale.edu

Paper prepared for delivery to the Annual Meeting of the Association for Public Policy Analysis and Management, Washington, DC, November 6, 2003.

Session on:
Environmental Regulatory Policy: Examining Alternative Approaches

DRAFT: DO NOT CITE WITHOUT PERMISSION
COMMENTS WELCOME

Developing a Multi-Disciplinary Evaluation of an Environmental Policy Innovation: Impacts of Forest Certification within the United States

I. Introduction

The arena of environmental policy analysis and management has been dramatically affected in the last 15 years by the emergence of an array of innovative policy initiatives that defy traditional categories of instrument choice (Gunningham and Sinclair 2002; Gunningham, Grabosky, and Sinclair 1998). From shared governance mechanisms where non-governmental actors have a responsibility for policy implementation (Cutler, Haufler, and Porter 1999; Haufler 2001; Salaman 2002), to voluntary programs initiated by industry associations (Rivera 2003, 2002; Kernaghan 1999; Prakash 1999), to firm level choices that increasingly take societal interests into account in their business practices (Prakash 2000; Cashore, Vertinsky, and Raizada 2001; Cashore and Vertinsky 2000; Vertinsky and Zietsma 1998; Zietsma and Vertinsky 1999-2001; Jennings and Zandbergen 1995), the world of policy analysis has been turned upside down. Many of these initiatives comprise, or are part of, a “basket” approach to instrument choice (Gunningham and Sinclair 2002; Gunningham, Grabosky, and Sinclair 1998) in which aspects of command and control, market-based incentives, voluntary programs, and educational approaches, as well as substantive and procedural approaches (Howlett 2000a), are all invoked simultaneously (and arguably work together) in an effort to produce behavioral change surrounding a given problem.

Recognition of these trends has led to two important research questions. Just what “mix” of instrument choices works best together? (Gunningham, Grabosky, and Sinclair 1998) and what are the ultimate impacts of these new arrangements in solving the problem for which they were created? That is, how *effective* are these new approaches? This paper is an effort to begin to address these questions by offering and justifying a protocol for comprehensively assessing

the impacts of policy innovations in reversing environmental deterioration. While the protocol could be applied to any policy instrument, we argue that these new and innovative “basket” policy instruments, which often do not get their policy making authority from traditional state-centered sovereign Westphalian authority, are especially in need of careful and comprehensive analysis. We develop this protocol inductively from our previous and ongoing research on forest certification (Cashore, Auld, and Newsom Forthcoming, 2003; Newsom et al. 2003; Auld, Cashore, and Newsom 2003; Cashore et al. 2003; Cashore, Auld, and Newsom 2002; Cashore 2002; Cashore and Lawson 2003), which represents a startling new non-state, market driven policy instrument in which standards are developed by non-governmental organizations. The market’s supply chain constitutes the arena in which evaluations by firms and actors about support occur, and hence where policy authority emerges.

We argue that such an effort is needed for three reasons. First, existing research within political science and comparative environmental policy often escapes or downplays concerns about “on the ground impacts” by describing and/or explaining policy decisions, rather than behavioral change. As a result, the literature either “black boxes” effectiveness, ignoring it altogether or assuming that environmental policy outputs *will* produce behavioral change in a direction that addresses the original policy problem; or the literature operationalizes effectiveness in ways that political scientists, rather than natural scientists, are able to measure (Young 1999).¹ While these treatments are important, less attention has been paid to actual changes in on the ground behavior and the implications of this behavioral change for improving the environmental conditions of the world’s biosphere, despite this being the ultimate concern for those who study environmental policy.

Second, sophisticated work within and outside of the policy sciences on applied policy analysis (Weimer and Vining 1999), cost-benefit analysis (Weitzman 1998; Metrick and Weitzman 1998), program evaluation (Rogers 2000), environmental assessments and environmental indicators (The Heinz Center 2002) tend to focus on key aspects of the problem, rather than explore how the key factors work together. Hence, applied policy analysis text-books tend to downplay, or treat as exogenous, such factors as the emergence of new institutional designs or the growing intersection of private and public authority. In these texts, policy instruments are treated as relatively static in their design and impacts, despite contrary empirical evidence on the structure and function of new innovative policy instruments.

Likewise sophisticated program evaluation work tends to explore only whether the program's original goals were met, rather than whether the original problem for which the program was created was ameliorated (Rogers 2000). And those who do study and explore actual conditions of the biosphere, (The Heinz Center 2002) (Esty 2002; Esty and Cornelius 2002) focus on describing indicators that they believe policy ought to address. This work has done an excellent job of improving our understanding of the problem, but careful attention to how policy has impacted these conditions, or how future policies might ameliorate them, as been given less attention.

Third, conducting broad scale analysis is not only a daunting task, there appears to be very little (career advancing) reward in attempting it. We show below that conducting such research means treating as *endogenous* factors that most analyses treat as *exogenous*. But not only does such an effort require a multi-disciplinary approach in what is still a disciplinary-focused world, attempting to do it quickly leads to the recognition that there are too many factors, and too many processes and interactions among variables, to develop an approach that

relies on quantitative measures alone, nor one that seeks absolute certainty. Recognition of this state of affairs can lead to considerable discomfort, causing proponents of interdisciplinary research to scurry back to the comfort of their own disciplines and expertise. We argue that these very disciplinary or ontological and epistemological biases against this type of work make attempts to do it even more important.² We also suspect that doing so might move policy making in directions that are more effective in reducing the impact of humans on naturally functioning ecosystems.

This paper proceeds in five parts. First, we briefly assess the challenges and hurdles in evaluating non-state market driven governance systems, which are emerging in a number of sectors including forestry. This step is important because the creation and support for these non-state market driven systems can be traced to the emergence of “liberal environmental” norms (Bernstein 2002, 2000) that coincided with environmental groups’ dissatisfaction with governmental approaches as doing too little, and industry dissatisfaction with government as being too controlling rather than with any *evidence* that these new instruments would be *more effective* in addressing global forest deterioration.

Second, the paper reviews existing directions within policy analysis and program evaluation literature, and argues that they present a starting point, rather than an endpoint, in developing a comprehensive evaluation framework. Third, turning to Bernstein and others (2000) and drawing on the previous section, we offer a five-step process for conducting comprehensive analysis of complex and dynamic “basket” policy instruments. Fourth, we illustrate the utility of this approach by applying it to our current evaluation of the impacts of forest certification in the conservation of biological diversity in the United States. We conclude by discussing the strengths and weaknesses of such an approach, and the need to do more

interdisciplinary work on refinement and reflection, even in the absence of institutional incentives to do so.

Failure to do this, we argue, often leads to emphasizing research programs of factors that are relatively easy to identify and/or quantify, while downplaying or ignoring equally important factors that are more difficult to measure (Patton 2002). We argue that the approach offered here would complement existing and outstanding work on economic cost/benefit analysis, econometrics modeling, and sophisticated efforts aimed at describing environmental conditions.

II. Emergence of Non-state Market Driven Governance Systems as “Basket” Policy Instruments

Arguably the best illustration of why a basket approach to policy instrument analysis is needed, and why on the ground effectiveness is at once complex but crucial to understand, is found in the emergence of environmental non-state market driven “certification programs” that have emerged domestically and globally in the last decade. These private governance systems are emerging in a number of business sectors globally and domestically (Table Two) including tourism, food production, mining and fisheries (Sasser 2002, 2003; Cashore, Auld, and Newsom Forthcoming: Chapter One; Conroy 2001), with arguably the most advanced case as we detail below, found in the forestry sector.

Cashore (2002) has identified four features (Table 1) of these systems that render them unique (Table 2). First, they have been created by non-governmental organizations and hence, do not rely state-centered Westphalian authority to adhere compliance to their rules. Second, and instead, the market’s supply chain provides the arena in which authority is granted.

Organizations and individuals along the supply chain, from forest owners to manufacturers to retailers to individual consumers, must make internal evaluations about whether to comply. As a

result, these instruments are particularly complicated because their source of authority is at once more complex, and less durable, than those that carry directly, or indirectly, state authority.

Fourth, compliance is verified usually in the form of an external audit undertaken by yet another organization.

These non-state market driven approaches have elements of all key types of policy instruments (Gunningham, Grabosky, and Sinclair 1998; Howlett 2000b). They are economic instruments in that firms along the market's supply chain evaluate whether they ought to support, or comply with, the certification system. In addition, environmental groups and their allies attempt to influence these evaluations by creating carrot-based green markets, or stick-based boycott campaigns. They are voluntary instruments in the sense that companies and individual forest owners cannot be fined or put in jail for failing to comply. They are informational/educational in that significant effort is placed on educating forest owners about forest management, and on encouraging "continual improvement". They have elements of "command and control" in that some of the certification systems' standards and rules that must be implemented required to obtain certification are *prescriptive*, requiring, for example, limited harvesting near streams and the creation of protected areas in the case of forest certification. In addition to these substantive foci, they also invoke and indeed emphasize procedural elements. There are clear procedures regarding which groups are to be involved in creating and developing standards, and firms being certified must undertake a number of procedural reforms, including developing management plans that include input from outside stakeholders, as well as developing firm-level procedures concerning an array of environmental and social values. The emergence of non-state market driven poses even more challenges to policy instrument analysis because, as Salamon (1981; 2002) notes, we cannot assume that what we know about cause and

effect within traditional public policy approaches can necessarily be transferred over to new shared or private governance systems (Salamon 1981; Salaman 2002).

Table 1: Key Conditions of Non-state Market Driven Governance

Role of the state	State does not use its sovereign authority to directly require adherence to rules
Role of the market	Products being regulated are demanded by purchasers further down the supply chain
Role of stakeholders and broader civil society	Authority is granted through an internal evaluative process
Enforcement	Compliance must be verified
Source (Cashore 2002)	

Table 2: Comparison of Emerging Non-state Market Driven Governance Systems Across Sectors

	Conditions of Non-state Market Driven Governance				
	State does not require direct adherence to rules	State does not control standard-setting process	Products are demanded by purchasers further down the supply chain	Authority is granted through evaluative process	Compliance is verified
Forestry	✓	✓	✓	✓	✓
Fisheries	✓	✓	✓	✓	✓
Coffee	✓	✓	✓	✓	✓
Organic Foods*	✓		✓	✓	✓
Ecotourism	✓	✓		✓	Usually
*refers to organic food certification in the US Source: (Cashore, Auld, and Newsom Forthcoming)					

III. Existing Policy Analysis Approaches

“...it is time we began to show by example that policy scientists do in fact have something constructive to offer the beleaguered policymakers. There is an obligation to alert society’s participants to impending problems; the policy scientist has a right to clarify and elucidate the competitiveness and complementarity of interests and objectives of those acting on behalf of society and of the society being acted upon; *we should not lose sight of the unique history and rich store of antecedent events leading up to the present and emerging future of various policy settings*; and finally, there is the unstinting

demand to behave as scientists, and to the greatest extent possible use, modify and create new tools, techniques, and explanations to understand as well as possible what impedes and supports the realization of better policies – purposive acts taken on society’s behalf” (Brewer 1974: 244) (*italics added*)

In the almost three decades that have transpired since Garry Brewer counseled us on where the future direction of the policy sciences ought to head³, there is no question that huge strides have been made in understanding and better exploring policy cycles (deLeon 1999), policy subsystems (Sabatier 1999b; Howlett and Ramesh 1995), the development of new and innovative instruments (Gunningham and Sinclair 2002; Gunningham, Grabosky, and Sinclair 1998; Howlett 2000b). There has also been significant effort in detailing how to conduct applied policy analysis (Weimer and Vining 1999; Clark 2002; Clark, Williard, and Cromley 2000)] economic cost-benefit analysis (Weitzman 1998; Metrick and Weitzman 1998) (Mendelsohn 2001) and program evaluation exploring whether a particular policy or problem achieved its goals. There is also significant natural science (The Heinz Center 2002) work identifying the staggering and devastating scale of global environmental deterioration. However, we seem to have made less progress in conceptualizing how the insights of this excellent work might be integrated to develop a comprehensive research agenda around a particular policy problem (For important exceptions, see Clark 2001, 1997). Instead, each of the above efforts have developed its own agenda and methods, with a strong bias towards identifying and measuring those factors that can be quantified (despite no evidence that non-quantifiable factors are less important than quantifiable ones), and a tendency to treat highly dynamic (but important) factors as exogenous.

These approaches render policy analysis easier, but arguably less relevant, than if they had conceptualized and incorporated the role of dynamic and non-quantifiable factors. The result has been that particular projects within the policy science framework have excelled, while efforts

to comprehensively address real-world problem-solving among policy scholars appear to have waned (Brunner 1991), This state of affairs has lead some scholars (Schön 1994) to claim that the policy science movement has largely failed. As Edward Parson has noted, “our ability to describe and understand the aggregate state of the natural environment, and its consequences for people, are limited by imperfect knowledge of natural systems, perceptual habits and biases, and disparate bases for valuing environmental attributes” (Parson 2002: 10-11).

Trends in Policy Analysis

The development of the policy sciences in the last 30 years has largely gone in the opposite direction than the multi-disciplinary and holistic approach envisioned by Brewer, with two quite distinct pathways emerging (Howlett and Ramesh 1995; Pal 1997). One pathway focuses more on explanatory work or theory building and another focuses on particular aspects of applied problem solving in which the end result is often a prescribed course of action. We review these pathways, and then offer a protocol designed to bring them back together and thus advance application of Brewer’s advice.

A. Explanatory Policy Analysis

The arena of explanatory policy analysis has taken enormous strides within and outside of political science. Public policy is now a sub-field within political science, and efforts at understanding institutions (North 1981) (Skocpol 1986, 1995) (Weaver and Rockman 1993), policy choices, (Pierson 2000, 1993) (Hacker 2001), policy networks and subsystems (Howlett 2002) (Marsh and Rhodes 1992b, 1992a) (Van Waarden 1992) (Sabatier 1991, 1999a) cite march asn Rhodes, Howlett, Hoberg, Sabatier] and local governance mechanisms (Ostrom 1998)

(Agrawal and Gibson 1999) and policy cycles (DeLeon 1999) (Anderson 1984), have resulted in much better understanding of the way in which political systems mediate conflict and produce policy choices. However, the emphasis within this literature on explaining policy choices and outcomes has meant that its practitioners rarely systematically assess the effectiveness of different policy instruments on the problems for which they were created.⁴ In general this literature treats policy outcomes as important end results (such as the shift from Keynesianism to Monetarism), without systematically addressing whether they address such goals as poverty reduction or environmental deterioration. Other studies that attempt to explain differences in the welfare state (Esping-Andersen 1989, 1990), including why some countries have universal health care systems and others do not (Hacker 2001), implicitly take as given the benefits of universal health care, paying less attention to prescribing a particular program. Or, in the case of pollution, political scientists tend to focus on those indicators that are measured (levels of particular pollutants escaping from a particular place of industrial production) (Harrison 1996, 1999), rather than on any detailed analysis of the impacts of such pollutants in affecting human health.⁵

While there are varying degrees of research designs, and many do not follow precisely Popperian methods, these works for the most part are efforts in which hypotheses or arguments are either tested or illustrated, hence following more or less accepted positivist approaches to gathering data (Appendix A)

B. Applied Policy Analysis

The other direction the policy sciences have followed is that of **applied policy analysis**, where researchers have undertaken efforts that usually result in the recommendation of a particular course of action. Here, the stages process first offered by Lasswell and modified by Brewer is more or less followed, even if particular steps might be given different labels

(Appendix B). A pattern has emerged, from a broad approach of those who more or less follow the Lasswell/Brewer approach to problem solving, to efforts involved in evaluating particular programs, to undertaking (legally mandated) environmental assessments of laws or programs, to comprehensive efforts that limit their efforts to developing and measuring indicators. We review these developments below and argue that a revised protocol for evaluating impacts should integrate, rather than disaggregate, these types of projects.

Comparing and Contrasting Different Policy Options

Many of those following Lasswell's and Brewer's advice have taken as their assumption the need to compare and contrast distinct courses of action over a given problem. Clearly following a linear stage approach, text books on applied policy analysis commonly refer to this technique (Weimer and Vining 1999) (Clark 2001). This work tends to be client driven, requires that values be explicitly identified and operationalized, and that impacts of each option be addressed according to different goals. While it is rare for journals to publish this type of work, it has influenced the approach used by many consultants, and generations of students educated in natural resource and environmental resource management schools. Under this treatment, policy instruments tend to be handled as relatively static, analysts are admonished to compare distinct policy instrument options (Weimer and Vining 1999), rather than explore whether they might intersect to produce more effective outcomes, and the goal of efficiency is usually given primacy over other goals, while issues of compliance are treated as political feasibility issues. The latter emphasis reflects the bias in exploring traditional governmental approaches, rather than new voluntary and market based instruments in which the very effort to gain support for them influences their (dynamic) form and function.

Program Evaluation⁶

Like policy analysis, program evaluation and the broader term of evaluative research stem as formal modes of inquiry from the policy sciences (Knapp and Kim 1998). Analysts, Leonard Rutman and George Mowbray, have defined program evaluation as “the use of scientific methods to measure the implementation and outcomes of programs, for decision making purposes” (Rutman and Mowbray, 1982: 12). Reflecting Lasswellian roots, David Nachmias describes program evaluation as, “the objective, systematic, empirical examination of the effects ongoing policies and public programs have on their targets in terms of the goals they are meant to achieve” (1979: 4).

As is the case in policy analysis, program evaluation has undergone a similar evolution – or revolution – with regard to questions of adherence to the scientific method. Earlier works (e.g., Suchman, 1967) emphasized that evaluation methods adhere to the basic logic and rules of the scientific method as closely as possible. Yet program evaluation scholars have come to acknowledge that methodological difficulties and lack of data often prevent the use of experimental or quantitative procedures. As a result, many scholars argue for the use of naturalistic or qualitative approaches on their own or alongside quantitative approaches (Haas and Springer, 1998; Patton, 1987; Herman, et al, 1987; Scriven, 1976). They also recognize, following Lasswell and Brewer, that choices about how to conduct program evaluation will necessarily impact results and hence purely scientific answers to problems are unattainable (Ingram and Mann, 1980; Howlett and Ramesh, 1995).⁷

While scholarship has emphasized the need for a variety of approaches that cope with imperfect information, applications of this approach have taken a different direction. One

technique that has emerged for working toward systematic and rigorous program evaluation entails constructing a logic model that depicts the connections between program inputs, activities and processes, outputs, immediate outcomes, and long-term impacts (Patton 2002). This is useful as it forced policy analysts to systematically think about the interactions among program elements within the broader environment in which policy is directed. Recognition of the dynamic nature of this process has led one step further beyond the logic model to what is called “a theory of change or theory of action model” which is explanatory versus descriptive and specifies hypothesized causal linkages (Patton 2002) (Rogers 2000). Development of a theory of change or theory of action model is one way in which program evaluation activities may identify causal contingencies happening within policies. In this way, some scholars view program evaluations as tests of policy theories, which contain all of the empirical and normative presuppositions and assumptions embodied by a program (Hoogerwerf, 1990).

To test a theory of change model, evaluators break down the model into a series of sub-objectives, putting the focus on small – and therefore more easily believable – observed causal relationships (Mohr, 1992). This technique briefly describes process tracing, a commonly used qualitative method (Collier 1991). To back up such inferences, evaluators often rely on triangulation in which several different data sources and collection methods are employed (Herman, 1987; Ragin, 1994; and Vedung, 1997). Multiplism, a special kind of triangulation, is enacted when evaluators employ both quantitative and qualitative methods in their analysis (Cook, 1985).

But what is different and important about this scholarship is that is decidedly comprehensive in its approach, first addressing the large questions and then moving to shed light on issues that influence the causal model. We argue that this is arguably the most important

contribution of existing scholarship on program evaluation, and any model designed to address effectiveness must likewise take a broad view.

Program Evaluation in Practice

Program evaluation has done much to advance applied policy evaluation, incorporating a range of factors and serving as a reminder that quantitative factors alone may provide misleading or inaccurate pictures. In practice however, difficulties in linking such analysis to the natural science world has often led program evaluation to assess goals identified either by the legislation creating the program, or by the program itself, rather than on the ground change in terms of reversing environmental deterioration.

Program evaluation most certainly has become ingrained in the culture of environmental policy (Knapp and Kim, 1998). The US Environmental Protection Agency, for instance, is required under the Government Performance and Results Act of 1993 to report on over 70 annual performance goals established around ten strategic goals (US EPA, 2003). The Agency faces many challenges in measuring actual outcomes from specific policies, much less across EPA programs in the aggregate (US EPA, 1994; US EPA, 2003). In light of the obstacles to measuring program outcomes, the EPA, like many other groups faced with evaluation requirements, many times resorts to quantifying program outputs, such as the number of new chemicals that are reviewed versus the hazards actually eliminated through such screening (US EPA, 2003).

(Environmental Protection Agency. United States 1997, 2003)As part of its reporting requirements, the EPA reports on performance concerning ten goals: clear air; clean water; safe food; preventing pollution and reducing risk; waste management; reduction of global and cross-

border risks; quality environmental information; sound science; credible deterrent and greater compliance; and effective management (EPA, 2003). The overall approach of the EPA's annual reporting efforts is that of an impacts evaluation, though lack of data and methodological limitations often result in the quantification of program outputs, rather than on impact in reversing environmental deterioration. For instance, in its Annual Report the EPA reveals how it tracked output goals, such as number of Superfund clean-ups performed or percent of water systems in compliance with standards, and does not link this performance to ultimate environmental outcomes. (A major exception are performance goals related to air quality (and in some cases, water quality), where required monitoring for some pollutants and the reliance on modeling mean that the Agency is able to provide estimates for reductions in illness and other outcomes).

The lack of data but need to quantify outputs has led to the EPA's annual reporting increasingly taking the direction of indicator assessment that we review below. There are debates in the Agency itself (e.g., see <http://www.epa.gov/evaluate/whatis.htm>) – about whether it is even doing program evaluation in many cases or rather simply identifying indicators and conducting performance measurement (detailed below). Certainly the EPA's overall approach does not lead to an integrated evaluation of outcomes across environmental media for an ecosystem view but rather treats water, air, etc. as discrete and independent areas. Improving this and other evaluation approaches and otherwise building evaluative capacity is a key area of concern for some in the EPA (see <http://www.epa.gov/evaluate/about.htm>). Efforts are now underway to develop and tracking comprehensive environmental indicators that would not be an end in itself, but would lead to improving the data available for program goal evaluations (<http://www.epa.gov/Envindicators/abouteei.htm>).

Environmental Assessments

There has been a plethora of work on environmental assessments largely owing to US legal requirements that such analysis take place. Environmental assessments in general are plagued by data and methodological problems, with the result that methods are usually focused on cost-benefit analysis of the current regulatory environment. The approach has taken to quantifying benefits and costs, ignoring to whether such an approach is possible (Tribe 1982), and producing evaluations of programs based on whether identified and quantifiable “benefits” outweigh identified and quantifiable “costs”.

Environmental assessments of the Clean Air Act illustrate this trend. Section 812 of the 1990 Clean Air Act Amendments (US EPA, 1997) require cost benefit analyses be undertaken. In 1991, the EPA released a report on the “The Benefits and Costs of the Clean Air Act, 1970 to 1990”. It addressed the overall question of how the health, welfare, ecological, and economic benefits of the Clean Air Act Programs compare with the costs of those programs between 1970 and 1990. It focuses in on the National Ambient Air Quality criteria pollutants: sulfur dioxide, nitrogen oxides, carbon monoxide, particulate matter, ozone, and lead. The study’s overall approach is one of a cost-effectiveness assessment. It is accomplished through cost-benefit analysis and ecological and econometric modeling. The study relies on models to estimate the following: costs and macroeconomic effects; air pollution; human health (e.g., particulate matter-related premature mortality) and visibility effects of criteria pollutants; ecological effects of criteria pollutants; and economic valuation of effects.

Thought the study does not address process, or implementation factors, it does develop and compare two regulatory scenarios. The “control scenario” reflects the actual conditions

resulting from the Clean Air Act Programs, and the “no-control” scenario reflects the expected conditions under a situation where air pollution control programs remained as they were before the passage of the 1970 Clean Air Act. In essence, the “no-control” scenario represents a hypothesized “baseline.”

While extremely useful in providing important information about the costs of environmental regulations, such approaches do not address other key factors that seem necessary for evaluating impacts including changes in political support, the arrangement of political coalitions, and emerging norms about the appropriate role of government and markets in regulating behavior. While such factors ought never to be ignored, this omission is particularly problematic concerning new policy innovations such as non-state market driven or other voluntary and educational approaches when issues of authority and support are far from given, and indeed as we show below, are directly related to the standards that emerge and the ultimate support that is granted.

Indicator Approaches

Increasing efforts at environmental analysis have used indicators approaches both to shed better light on the nature of the policy problem but also as means in themselves – as benign educational approaches that it is hoped, will encourage governments, firms, and individuals to act in different ways once information about their behavior, or the state of the natural environment over which they have policy authority, is better understood. These efforts are asserted to have independent impacts in influencing behavior and hence reversing or addressing environmental and social deterioration (Florini 2003), though little systematic work has been done to evaluate these claims.

One of the most expansive and exhaustive efforts in this regard was undertaken by the H. John Heinz III Center for Science, Economics, and the Environment (The Heinz Center 2002), which sought to lay out a design for periodic reporting on the condition and use of the ecosystems in the US. The effort developed one of the most innovative and comprehensive studies on the state of the US environment, involving close to 150 individuals and organizations at the forefront of environmental science, advocacy, and public policy over a period of five years.⁸ Its objective was to provide the inputs needed for evaluations by initiating a “periodic, comprehensive, and reliable compilation of essential information about the overall state of the nation’s environment” (The Heinz Center 2002: 3). The study’s overall approach is to use indicators, and it consciously makes no effort to connect the data back to potential explanatory factors.⁹ Ten “core national indicators” are also reported on to provide an overall, cross-media picture of national ecosystem conditions and usage. The Heinz report and work is extremely valuable in detailing knowledge about the state of a policy problem – an invaluable effort for anyone interesting in addressing environmental deterioration. Indeed, such efforts should be

placed at the beginning of any evaluation, so that the problem itself is understood as best as possible. But what it explicitly shies away from is any effort to develop a causal model about what affects the indicators, let alone any attempt to measure such mediating factors.

Despite these limitations, the use of indicators has increased, and indeed indicators are sometimes used as ends themselves, rather than as methods for identifying policy problems. Indicator development has become extremely popular in the forest sector, where intergovernmental failures to achieve a global forest convention refocused governmental efforts toward defining a common system of indicators for sustainable forest management. Two key processes include the Montreal and Helsinki approaches to developing indicators of sustainable forest management. The intergovernmental processes have involved an array of stakeholders to identify what ought to be considered when measuring sustainable forestry, with the presumption that work aimed at assessing the state of the forests can be more systematic, comprehensive, and transparent. The complex nature of forest ecosystem structures and the influence of communities and business that rely on them have made this process as politically influenced as any traditional policy instrument. Interest groups and professional foresters actively engage in these efforts in order to lobby for their own preferred emphasis on economic, social, or environmental indicators. The result has been the development of sustainable forest management indicators that are the product of often heated debates and disputes among parties at the table about what ought to be measured, and what ought to be emphasized. Some advocates of indicator development go further, claiming that objective science can provide an answer as to the appropriate “balance” between environmental, social, and economic indicators (Oliver 2003).

IV. Towards a Protocol for Evaluation Research

The current work on applied policy evaluation has gone in important directions and produced important knowledge, but in practice in most cases it has also moved decidedly away from the big picture and systematic approaches recommended by the founding fathers of the policy sciences, and more into narrow, and decidedly more benign approaches to instrument choice, with little information about their actual impacts. Important exceptions to these broad trends reveal the usefulness in building broader and more systematic approaches. For example, the US congress commissioned a two year National Academy of Public Administration (NAPA) report (2000) designed to develop lessons learned from environmental policy innovations. Over the course of the two-year effort, different scholars were asked to perform assessments of innovative policy programs started by the EPA and its partners. The programs studied ranged from state-based cooperative environmental agreements to air pollutant trading systems, and while specific methods differed, all comprise single case studies.¹⁰

A clear effort was made to integrate the findings of individual studies by exploring the impact of the instrument in changes to environmental conditions where possible, and to then identifying brief futures scenarios in which the differing approaches are combined in different ways leading to varying larger governance implications. Research focused on identifying support for new initiatives among firms, explaining differences, and then where possible, linking support to quantifiable outcomes such as pollution emissions. This effort recognized the need to aggregate different efforts involved in policy evaluation noted above, and provides evidence about the need to conduct more systematic research.

Similarly the EPA appears to have recognized the importance of an integrated approach, commissioning an independent review of one of its largest innovations programs, the Common

Sense Initiative (CSI). CSI consisted of multi-stakeholder pilot project development in six industrial sectors for environmental policy that is “cleaner, cheaper, smarter.” The evaluation report (Kerr et al. 1999) was designed to assess the progress of and lessons learned from the four year project, and involved a range of qualitative, primary, and secondary data collection to address a range of factors.¹¹ Instead of providing concrete answers to narrow questions, the study acknowledged that CSI was preliminary in nature, and that future work would be required to understand better the impacts of the program and its effectiveness in addressing environmental deterioration. By starting with broad questions, rather than focusing efforts on quantifiable factors, the study was arguably well positioned to identify future research needs, and to reflect on what future impacts might be, than if it had only focused on more narrow but easily quantifiable factors. These broad approaches to policy evaluation illustrate the need and benefits in developing evaluations that depart slightly from pure scientific methods, in order to produce more comprehensive and rigorous evaluations.

Applied Forward Reasoning: Overcoming Physics Envy

What is the way to proceed then with understanding the potential effectiveness of a particular policy instrument in addressing or reversing environmental deterioration? We argue that any evaluation work must begin with the difficult question of effectiveness, and then develop research questions around this. We turn to Bernstein and others’ (Bernstein et al. 2000) effort in explanatory research, in which they asserted that the sub-discipline of international relations had become so preoccupied with following positivist Popperian and (usually) quantitative techniques aimed at precision, that the big questions of our day, such as the causes of war, violence and environmental degradation, were never explored systematically. We argue

the same trend is true in applied policy analysis, and that the lessons from Bernstein and other (2000) can be used in revising steps in undertaking applied policy evaluation.

Towards a Systemic Evaluation Framework

An alternative to simply counting outputs that might miss complex dynamic processes and avoid key indicators that cannot be quantified is addressed by Bernstein and colleagues (2000). These authors suggest that the nature, goals, and criteria for judging social scientific enterprise must be rethought if the results are to be more helpful in understanding the complex questions of the real world and in aiding decision-making in the face of uncertainty. The first step in their approach is to “overcom[e] physics envy,” (48) or appreciate key inherent distinctions between the physical and social sciences. These differences include the fact that in social science, we study open, non-linear systems about which the development of predictive theory is virtually infeasible. As such, rather than attempt point predictions, the policy-relevant social scientific pursuit is more appropriately intended for the “identification and connection of chains of contingencies that could shape the future” (53).

Bernstein and others (2000) put forward a strategy of research to suit this purpose they identify as “forward reasoning”. The strategy is centered around the development and use of scenarios, or narratives that “map a set of causes and trends in future time” (Bernstein et al. 2000: 53):

Scenarios are impressionistic pictures that build on different combinations of causal variables that may also take on different values in different scenarios. Thus it is possible to construct scenarios without pre-existing firm proof of theoretical claims that meet strict positivist standards. The foundation for scenarios is made up of provisional assumptions and casual claims. These become the subject of revision and updating more than testing. (54)

In all, the forward reasoning process has seven steps. The first is identification of driving forces, which are the causal elements surrounding a problem and can be the basis – in various combinations – for different chains of connections and outcomes. The second step is specification of predetermined elements, and the third is identification of critical uncertainties. Fourth is the actual development of the scenarios with plot lines that describe how driving forces might interact with the predetermined elements and different combinations of critical uncertainties. The fifth stage is assessment of early indicators of the scenario, and the sixth is the consideration of scenario implications. In the final step, low probability events or actions, known as wild cards, are conceived of and evaluated with respect to their potential effects on the chains of logic defining the scenario.

Alternative Analytical Approach: Steps In Applied Forward Reasoning

There is every reason to believe that the lessons from Bernstein can not only be applied to policy evaluation, but that its application might serve to reunite the explanatory and applied paths of the policy sciences, redirecting and refocusing efforts toward Brewer’s original mission. We will call our approach, “steps in applied forward reasoning” (Appendix B). While we apply these steps in a linear fashion, we recognize that working within this process will necessarily involve an iterative process.

- I. **First Step: Identify Problem**
 - a. **Specify problem clearly.**

The first step is to carefully understand and identify the extent and nature of the problem. Natural science indicators should be developed, turning to both quantitative and qualitative assessments. It is important not to collapse different dependent variables into one, such as

identifying environmental and social indicators, which are bound to have different explanatory causes and different impacts from traditional and innovative policy instruments.

- II. **Second step:** Specify solution being proposed
 - a. **Specify policy instrument** being assessed.

This step involves carefully specifying the proposed policy innovation and the type of approach it is invoking to address the problem. What is the source of authority? What approach to environmental problem solving does the instrument take? Coercive? Voluntary? How does it compare to existing approaches? At this stage it is useful to focus on a single policy instrument at a time, since each instrument will influence and affect outcomes differently.

- III. **Third Step:** develop causal model of mediating factors
 - a. **Identify (explanatory) factors**

This step involves conceptualizing the universe of key factors that will affect directly or indirectly, whether individuals and firms will comply with the proposed policy innovation. That is, what factors are likely to shape/change behavior modified by policy instrument?. How do individual and organizational evaluations matter? Do market incentives make a difference? Are values and norms key in influencing on the ground change in the problem?

- b. **Explicitly trace a preliminary causal model** on how key factors interact to produce on the ground change.

Develop a model for identifying how key factors that determine support interact. Do they change over time? How might they change over time?

- IV. **Fourth Step:** Develop research program

Develop a research program built around enhancing knowledge of the component parts. Given the dynamic nature of a policy instrument, time constraints, and other factors, what are key things policy makers need to know? Can they be quantified? Does quantifying them contribute to broader knowledge? Identify those factors that can probably never be quantified. What would happen to the environmental problem if there was universal acceptance of the policy innovation? Would the problem be ameliorated? If so, to what extent?

- a. **Identify Impacts (dependent variable) if complied with**

What does the policy innovation specify must be done? What do the standards, policy and rules say? How do they relate to the original problem? Are their specific components? What does existing work in natural science tell us about potential impacts on behavioral change? What research would better uncover this relationship, given the causal factors dealt with above?

V. Fifth Step: Plausible Scenarios

- a. Identify futures plausible scenarios that emerge

Identify what appear to be the key types of future scenarios that might exist, based on the research project. Undertake comparisons of policy alternatives based on plausible scenarios. Identify the most promising new arenas of future research to improve forward reasoning, and those areas that are less important. Leave comparisons of other policy innovations for this stage, in order to better consider future scenarios that might not have been evident in the beginning of the research project.

VI. Sixth and Seventh Steps: Early Indicators and Implications

- a. Identify key indicators about the scenarios and potential implications.

Applying Applied Forward Reading: Impacts of Forest Certification within the United States

The case of forest certification and its impacts in the United States fits well with the applied forward reasoning approach offered here. Indeed, the story below indicates that the complexity of the policy situation under investigation precludes a strictly experimental approach and instead requires a mixture of quantitative and qualitative methods to be fully appreciated. Indeed, the analyses conducted thus far of the instrument of forest certification provide an example of how overly narrow definitions of program effectiveness have been used to make up for lack of data or feasible research methodologies. Eco-labeling – i.e., the general term covering certification – has been the subject of many studies, but most of the studies have examined projected (i.e., attitudinal) responses from consumers rather than the actual price premiums or other market advantages companies realize from eco-labels (Mayer, Lewis, and Scammon, 2001; Sasser, 2002), much less the environmental or social improvements related to the functioning of eco-labeling systems. Studies focusing on the actual market behavior relating

to eco-labeling, or certification, within the forest products industry are virtually non-existent (Sasser, 2002).¹²

Despite all of the attention the FSC and forest certification in general have garnered over the past decade, the effects attributed to them are almost without exception anecdotal rather than empirical (Sasser, 2002). In our view, it is no accident that systematic, rigorous analysis of the impacts of forest certification is uncommon given its nature as a complex, basket policy approach composed of market-based, educational, and other aspects.¹³

The Evaluation Project

We originally began this evaluation project because we received a grant from the Doris Duke Charitable Trust, which included evaluating the effects of the Forest Stewardship Council on the conservation of biodiversity in the United States. We will also include in our illustration of the utility of the framework by making reference to other related research conducted on forest certification in the last four years by Cashore, Auld, Newsom, and Lawson. We do not claim that our research followed directly the exact linear path described above, but that such an approach, used as a guide, presents a useful approach designed to advance applied problem solving, especially regarding the increasing use of voluntary, market-based, and non-state dynamic “basket” policy instruments.

This policy problem represents a challenging, real-world question that pushes the bounds of policy analysis methods. For example, how do we assess the impacts when they are long-term, but FSC certification has been implemented for less than a decade? How do we assess the impacts when the policy approach is still undergoing change in all of its different basket aspects?

Within our analysis, can we isolate the FSC certification program from other ongoing forest policy approaches?

We make use of this research task and its challenges to explore the potential of the applied forward reasoning approach. We briefly detail the research questions that arise from such a framework, the results to date of applying the research, and what research is ongoing. In the final section, we reflect on the development and utility of this approach in the context of the current research at hand and the fields of policy analysis and program evaluation in general. ¹⁴

1) First Step: Identify Problem Specify problem clearly.

The client has identified the desire to improve the conservation of (forest) biological diversity in the United States. While biodiversity is itself a complex phenomenon, it is a much clearer concept than sustainable forestry or sustainable forestry management as they have economic and social components as well. While all problems are important, it would be a mistake for any analysis to assume that a particular policy instrument would behave the same way towards social as well as environmental values. Having focused the dependent variable on the conservation of biological diversity, key members of our research team have developed key indicators regarding forest biodiversity conservation (See Appendix C). With these indicators, we are now in a position to carefully assess their condition and whether and how forest certification might influence their conservation. Speaking in broad terms, much is known generally about the state of key indicators of biodiversity conservation. Increasing forest fragmentation, logging in old growth forests, road building, harvesting in riparian zones, conversion of forest lands to other uses has placed considerable stress on forest ecosystems in the

United States. The list of forest-dependent species that are endangered or threatened is increasing in the United States, as are threats to the structure and function of forest ecosystems.

2) **Second step:** Specify solution being proposed

The solution being proposed is referred to as forest certification, and represents, as discussed above, a non-state market driven approach to addressing sustainable forestry. In the case of FSC certification, a hierarchy of ten principles, over fifty criteria, and at least three times as many country or region-specific indicators put forth the performance and process-based forest practice standards that encapsulate the FSC vision of well-managed forests. Undergoing assessment against these standards by an FSC-accredited third party is also part of the compliance package. In our analysis, we need to compare the requirements (i.e., the indicators) across the specific regions being examined within our study in order to find a way to look at certification impacts in an overall US context. Because FSC certification represents a governance system and so is more than a collection of forest practice standards, policies and rules about how the entire system is designed to work are also critical to understanding the end results of certification. In some cases, the connection between a FSC policy and the problem of biodiversity conservation will be plain (e.g., in the case of forest practice standards about riparian zone protection). In others, the link will be less direct but may still hold significance (e.g., in the case of decision-making procedures). Support for the FSC comes not from governments, but from the market's supply chain where firms who purchase products are coerced or cajoled into supporting the FSC through positive recognition, or negative boycott campaigns (Sasser 2003; Gereffi, Garcia-Johnson, and Sasser 2001).

Forest certification has been embraced by non-governmental organizations and other members of the forest policy community as a more efficient and effective route than traditional governmental approaches to address global deterioration of the worlds’ forests. The FSC was created in 1993 following intense dissatisfaction on the part of international environmental groups with international and domestic government-initiated efforts to address global forest deterioration. The US FSC national initiative, which has been in place since the 1996, is charged with overseeing the development of the program in this country (Cashore, Auld, and Newsom Forthcoming). The creation of the FSC, as we reveal below, has resulted in the development of a number of alternative certification programs initiated by forest owners or industrial forest companies. Industry has embraced the voluntary and market based approaches of these systems, offering their own versions (Table 3) rather than directly critiquing the policy instrument itself.

Table. 3, Conception of forest sector non-state market driven certification governance systems

	Conception One (Original FSC vision)	Conception Two (original SFI vision)
Who participates in rule making	Environmental and social interests participate with business interests	Business-led
Rules -- substantive	Non-discretionary	Discretionary-flexible
Rules -- procedural	To facilitate implementation of substantive rules	End in itself (belief that procedural rules will result in decreased environmental impact)
Policy Scope	Broad (includes rules on labor and indigenous rights and wide ranging environmental impacts)	Narrower (forestry management rules and continual improvement)

Source: (Cashore, Auld, and Newsom Forthcoming)

Table 1.4, Comparison of FSC and FSC competitor programs

	FSC	SFI
Origin	Environmental groups, socially concerned retailers	Industry

Types of Standards: Performance or Systems-based	Performance emphasis	Combination
Territorial focus	International	National/bi-national
Third party verification of individual ownerships	Required	Optional
Chain of custody	Yes	No
Eco-label or logo	Label and Logo	Logo, label emerging

Source: (Cashore, Auld, and Newsom Forthcoming)

3) Third Step: Construction of a Causal Model of Mediating Factors

In the third step, we utilize the framework of mediating factors to develop a research agenda.

The research agenda is built around the goal of enhancing knowledge about the mediating factors and the interactions among them. Specific questions about the FSC program practices and outcomes also serve to guide the development of the research agenda. Time and other resource constraints must be kept in mind in order to determine those components of the analysis that are most essential. In accord with the real-world orientation advocated by Bernstein, et al and by Lasswell and his followers, the determination of what is essential is made with reference to the most pressing information needs of the policy-makers in this case, the foundation clients. In the case of the analysis at hand, resources are not available to construct research methodologies that focus on the supply chain factors, the support from environmental and social NGOs, and the procedures of verification of certification standard compliance.

Based on the clients’ information needs and an initial review of the framework of mediating factors, a series of goals, objectives, and criteria are set forth as the metrics against which FSC program activities and outcomes will be assessed.¹⁵ These are the specific questions referred to in the paragraph above. We have developed four research projects that address the

various goals, objectives, and criteria and correspond to different sets of the mediating factors. In the following paragraphs, we briefly review these research projects and explain how they relate to the framework of mediating factors. For each project, we comment not only on the methodologies to be used but the nature of the data to be collected and how it may fit within the overall analysis.

Given the nature of the certification instrument we need to focus both on the actual standards and how auditing occurs, but also support in the market place for such a system and whether and how forest owners might be encouraged to support forest certification. And given the complex governing system, we need to know how this system develops and modifies standards, and whether and how the FSC is able to gain support by changing or adapting to forest owners or supply chain interests. Appendix D identifies our efforts at developing a framework of mediating factors that we have identified inductively through various research efforts (Lawson and Cashore 2001; Bernstein and Cashore 2000; Auld, Cashore, and Newsom 2001; Cashore, Auld, and Newsom Forthcoming). We have also benefited from others' research (Meidinger 1999, 2001; Elliott and Schlaepfer 2001).

The main components of the program logic model are the policy-making institutions of the FSC (e.g., the General Assembly and the various standard-setting groups); the forest practice standards generated (i.e., the principles, criteria, and indicators); the verification procedures for ensuring compliance with the standards and the other policies of the FSC; and the biodiversity conservation and other impacts that are expected to result from implementation of FSC certification. To this beginning, we start adding other key factors that influence the elements of the FSC program.

One is environmental and social NGO support. Another is the support that may be expressed by actors (manufacturers, retailers, and consumers) along the supply chain. Both NGO and supply chain support may exert pressure on the FSC policy-making institutions to produce more stringent forest practices standards. Another key mediating factor is the support that may be given by various types of forest companies or landowners. These stakeholders may increase their support for – or involvement in – the forest certification program if the program produces policies they feel are reasonable, thus potentially leading to pressure on FSC policy-making bodies to lower standards. Thus, in the framework the arrow goes both ways between FSC institutions and forest company and landowner support. Incentives, such as price premiums or shields from market campaigns, may tip the scales in favour of support for certification. At the same time, disincentives, such as the cost of certification or the increased scrutiny it may bring, may have a negative influence on companies' decisions about whether to take part.

Beyond the mediating factors themselves, there are other aspects of the framework worth emphasizing. One, hinted at above, is the fact that the interaction between several factors is two-way. In other words, factor X may affect factor Y independent of factor Y affecting factor X. Related to this aspect is the fact that the factors do not interact in any clockwork or systematic fashion. Supply chain support, for instance, may impact the policy institutions of the FSC during policy formulation, decision-making, implementation, or evaluation. Again, the policy-making process in either the public or private sphere is not comprised of truly discrete steps.

The framework is constructed in a way to demonstrate that the forest certification system does not remain static. As one factor affects another, that factor itself may be simultaneously impacted by yet another factor. We make use of the framework as a tool to begin to explore

potential combinations of changes to the system, and as we discuss below, how these changes might ultimately affect policies and outcomes.

4) Fourth Step: Develop Research Program

The next stage involves developing independent descriptive and explanatory research projects that will shed light on one or more factors identified above (Attachment A): forest owner support for FSC; institutional governance; supply chain support; economic incentives and disincentives; standards content, auditing; and the conservation of biological diversity (for simplicity, we have collapsed ecological function with biodiversity). We now review key projects we have developed or are developing that address these issues.

Mail Surveys Regarding Forest Owner Support

Early on in our research we conducted mail surveys to understand better non-industrial (Newsom et al. 2003) and industrial (Auld, Cashore, and Newsom 2003) forest owner support for the FSC, SFI and other voluntary sustainable forestry programs. This was collected in the form of quantitative data and provided the first snapshot into (weak) support among industrial forest companies for the FSC, and relatively stronger support for the SFI. We also found that non-industrial private forest owners preferred voluntary programs initiated by their own associations, and were most suspicious of the FSC. We also applied mail surveys in Germany, and then compared our data with similar efforts in Canada to better understand whether export-dependent companies preferred the FSC or system based and voluntary ISO 14001 approaches (Cashore et al. 2003).

Conceptual Framework on Types of Support

We (Cashore 2002) simultaneously developed a conceptual framework for understanding forest owner support that distinguished support based on economic reasons (pragmatic legitimacy) versus support granted owing to established values (moral legitimacy) or because support was seen as an accepted every day practice (cognitive legitimacy). Such a framework also permitted us to analyse dynamic parts of the model above because it identified “legitimacy achievement” strategies in which the FSC might conform to forest companies by reducing its standards; convert forest companies through the use of market-based campaigns; or inform likeminded audiences that they exist through advertising.

Comparative Case Studies

We also undertook comparative regional and country level case studies to understand why it was that in some countries and regions such as Sweden, British Columbia, Canada, and the Maritimes the FSC appeared to be gaining initial support while demand in the US was so weak (Cashore, Auld, and Newsom 2003) (Cashore and Lawson 2003; Lawson and Cashore 2001). Through inductive comparisons, we identified seven features of the forest sector that if present, facilitated FSC efforts to gain support and if not present, facilitated competitor programs such as the SFI. This largely historical and qualitative work permitted us to quickly learn what the key mediating factors were in influencing general trends in industry support in the United States. What this study did not do was to explain important differences within firm behavior in the United States, which, as we detail below, we are now in the process of analyzing.

Understory and Overstory Floristics Study

The floristics study concentrates on what actually happens on the ground after a forest owner is FSC certified. The purpose of the research component is to better understand the effects of forest practice guidance provided by regional FSC standards on the floristic diversity patterns at the patch, stand, and landscape levels. The study has this focus because maintenance of biodiversity is directly related to the demography and conditions of both timber (e.g., overstory species) and non-commodity forest species, including herbaceous understory plants, and their treatment within silvicultural practices.

In this research component, we will complete both secondary and primary data collection and analysis. For both understory and overstory aspects, the primary research includes profiling sampling areas, calculating various measures, and comparing results against the types of forest practices enacted within the different areas.¹⁶ Statistical methods will be utilized in the analysis of findings. Because of resource constraints, the primary research will be performed all within the Northeast of the US. To see if it is possible to link the findings about the Northeast to other areas within the US, the quasi-experimental approaches of the floristics primary research will be supplemented by the systematically conducted literature review and by shadow controls comprised of experts familiar with forest practices and conditions in other parts of the country. Shadow controls are an approach often used in qualitative program evaluation techniques (Rossi and Freeman, 1985; Vedung, 1997).

In this way, quantitative data will be supplemented with qualitative data. Qualitative data is appropriate for this research component because it will not be possible to produce quantitative estimates of environmental outcomes of FSC certification (e.g., 0.5 percent greater protection of species riches in the Northeast) based on the findings. As one of the primary sources of

information on environmental outcomes within the FSC analysis, we will explore ways in which the findings may be compared against findings from other research components. For instance, observations on feasibility and other aspects of the FSC forest practices garnered from the floristics secondary and primary research will be triangulated against data collected on such matters from companies within the firm level study and possibly from the regional standards assessment.

Landscape Management System (LMS) Study

The Landscape Management System (LMS) study also concentrates on the mediating factor of environmental outcomes of FSC certification. LMS is a computer forecasting tool that uses forest management plan information to predict outcomes from different forest practices. In our analysis, we use LMS to evaluate how effective FSC certification is in achieving the conservation of biodiversity and ecosystem functions within case study sites throughout the US. In brief, the procedure in this research component is to forecast the results of forest practices in FSC-certified as well as uncertified forests and compare these results. The comparisons will be made between the pre-certification and post-certification forest practices of a given landscape as well as between the forest practices of different FSC-certified and non-certified ownerships.

Use of a computer forecasting tool like LMS does much to address the data problems associated with environmental assessments discussed earlier in the paper. LMS helps us cope with the lack of longitudinal studies of forest practices and also allows results about case studies to be extrapolated to a larger level. By varying inputs into the LMS calculations or the calculations themselves, we can run the same analysis under a variety of geographical and other conditions. By varying the model and running it multiple times, we can aggregate results to say

something about the effects of FSC certification nationwide. LMS translates forest practice information into numerical equations and generates quantitative results.

LMS computer modeling takes the place of a truly experimental design, but it is important to note the qualitative aspects of the method. These include the use of shadow control expert judgment in deciding what is important to include within LMS, how to write the functions, and how to interpret the output. Given all of the uncertainties and simplification inherent in the modeling, we most likely will not emphasize quantitative point or range estimates but rather general trends revealed by LMS. By concentrating on general trends, it will be easier to attempt to connect the outcome results to other findings within the overall research effort. Moreover, the expert judgment used to program LMS can help us in formulating other areas of research, such as determining what parts of the forest practice standards should be focused on within the prescriptiveness assessment discussed below.

Preconditions and Conditions Analysis

In order to explore direct impacts of FSC certification on forest company behavior, we have launched a cooperative research agreement with Smartwood (a leading FSC auditor) in which they will analyze and detail the conditions auditors placed on all companies who went through FSC audits. The benefit of this data is that it identifies what changes companies must take before they can be certified, or in order to maintain their certification. The analysis focuses both on required procedural changes, as well as on the ground forest practices.

Firm Level Study

A project with collaborators at Duke University (Erika Sasser), University of Washington (Aseem Prakash) and Yale (Cashore and Auld) is designed to understand better the types of

choices forest companies make about forest certification. The firm level study corresponds to the following factors within the framework: forest company and landowner support, incentives for certification, and disincentives for certification. It seeks to explore and trace the factors affecting the certification choices of US forest companies, including whether they choose to support the FSC, or third party auditing under the SFI, as the industry alternative to the FSC. Explanatory factors include assessing the influence of market-based boycott campaigns and the role of associational systems in encouraging support for certification.

Once companies have been chosen according to our specific case selection criteria, secondary and primary data collection commences. We draw on company and industry reports, press, and other sources to construct company profiles, which help us to understand the environment in which each firm operates and the pressures imposed by outside sources (e.g., the mediating factors of certification incentives and disincentives). We conduct semi-structured personal interviews with company officials to determine the factors (used differently here) companies believe to be the most important in shaping their certification choices.

US FSC Institutional Study

The US FSC institutional study corresponds to the FSC policy-making institution factor within the framework as well as to the forest practice standards factor. This study seeks to explore those institutional aspects – at the regional, national, and international levels – that most affect the way in which the US FSC functions in the FSC governance system and works to impact forest practices within the US. Because the subject of the analysis is on the FSC in the US, the institutional component focuses in on the federalist aspects of the overall FSC system.

Also, because a primary activity of the US FSC thus far has been the regional standard-setting processes, they comprise a large part of the institutional study.

The basic task of the study component on the regional standard-setting processes is a comparison of the processes within four case study regions (the Northeast, Southeast, Pacific Coast, and the Great Lakes-Central Hardwoods States) and the standards that resulted from these processes. The methodology for the regional comparison includes primary and secondary data collection on the standard-setting processes in each of the regions and an in-depth analysis of each set of standards. The data collection will be performed with an eye toward identifying factors that might influence the standard-setting processes and the characteristics of the standards that result. As part of the analysis, we are performing an assessment of the standards in terms of their prescriptiveness with regard to some key issue areas.¹⁷ The results from the assessment will be cross-walked against the factors identified in the broader research done on the regional processes. We will make use of process tracing to analyze the interaction of variables to determine if the content of the standards can be connected back to standard-setting process characteristics.

We will then attempt to link the regional standard assessment back to larger questions about the FSC institutional structure. In essence, the regional standard assessment will be treated as an extensive case study of the federalism inherent in the structure. We will trace changes made to regional standards as part of the process to harmonize the US regional standards. We will also examine the influence of the FSC Secretariat on the regional standards during its review of them and also during the international endorsement process.

The regional standards assessment will make use of numerical systems as an aid to processing data and performing the comparison across the standards. The numerical scores

granted to standards will not, however, be emphasized in the final results because imprecision (e.g., prescribed lack of specific endpoints for forest practices) and a lack of consistency (i.e., in how they are constructed) across the sets of standards means that it will not be appropriate to extrapolate relative degrees of potential impact from the prescriptiveness assessment. That said, a final goal of the institutional study overall will be to determine if findings about the standard-setting processes, the standards produced, or other aspects of the FSC institutional structure can shed light on other factors within the framework, such as landowner and forest company support or environmental outcomes of certification.¹⁸

Step Five: Future Scenarios (*Implementation of Forward Reasoning*)

From the preceding discussion, it is apparent that there are many levels of analysis that go into our analytical approach. The first level of analysis is performed according to each individual project's methodology. Then, the findings from each project are cross-walked against one another to search for potential triangulation or correlation. This step may reduce uncertainty about findings or emphasize particular findings as especially significant. The final level of analysis brings in the techniques of forward reasoning to ensure that we do not treat the forest certification system as static.

In this final level of analysis, we compare key baskets of policy instruments that have emerged or may come into existence against the pre-specified goals, objectives, and criteria introduced in the previous step of our approach. The differing baskets of instruments correspond to the scenarios concept advocated by Bernstein, et al (2000). In our case, the key policy baskets are comprised of differing combinations of non-state certification programs and public policy or public/private approaches.

Plausible Scenarios: Identify Plausible Futures Scenarios that Emerge

While our research is not yet complete, the approach outlined here has allowed us to identify and rule out certain future scenarios over forest certification in the United States. What can be ruled out is a future in which all forest-lands are FSC certified. Non-industrial forest owners have, by a vast majority, either expressed no interest in forest certification at all, or a preference for voluntary landowner association programs such as Tree Farm and Green Tag. We can also rule out a future in which most industrial forest companies support the FSC, since the vast majority of industrial forest companies have opted for the industry alternative, the Sustainable Forestry Initiative, which has developed what the companies believe are more flexible and appropriate standards for industrial forestry (Cashore, Auld, and Newsom Forthcoming). The SFI, in turn, has a written agreement with Tree Farm that it will support that program for industrial forest lands. This means that there are essentially three likely scenarios for the future of forest certification in the United States (Following Bernstein and others, we also identify here early indicators for each scenario and what would be needed to assess its implications for biodiversity conservation in the United States):

a. Scenario One: the FSC Maintains Market Support

Under this scenario the FSC maintains some supply chain and pockets of forest owner support, as well the reputation for being the “greenest” certification option. In this case those limited number of companies who wanted to be seen as green leaders and, who did not mind breaking with associational pressure, would support the FSC. Under this scenario, the FSC remains a viable though minor threat to the SFI, with the result that the SFI either maintains, or increases its standards governing on the ground forest management.

Under this scenario the FSC’s biggest contribution is in the emergence and evolution of the SFI as a certification program.¹⁹ Any research into the effect of the FSC on biodiversity conservation would therefore have to assess the impacts of the SFI on firm forest management changes, and the impacts of those changes on biodiversity indicators. Hence, a research program directed toward better understanding the impacts of SFI standards on firm-level forests practices, and then the effect of these practices in biodiversity conservation, would need to be undertaken.

b. Scenario Two: FSC Dies

Under this scenario the FSC is seen as so weak by its funders and supporters that support is withdrawn, leaving the SFI as the only viable certification system in the United States. If this occurred, our existing research would indicate that special attention be placed on understanding industrial support for the SFI, and whether standards development might be relaxed, give the lack of pressure from the FSC. Especially important would be to trace whether the SFI is so accepted it has gained “cognitive legitimacy” or whether a certain degree of pressure from outside is needed to maintain industry resolve.

c. Scenario Three: Certification Dies

Under this scenario both the FSC and SFI are deemed by environmental groups and the public as being ineffective in addressing species loss and biodiversity conservation in the United States. If this were to happen, the question would be then whether more innovative voluntary instruments would be attempted, whether there would be a move to go toward additional traditional governmental regulation, or whether an approach using public and private authority

might emerge. Under this scenario, research and analysis would need to be focused on what other potential mix of instruments might work, and whether innovations in other sectors or countries might provide additional policy options.

While unlikely, a wild card, or low probability event, that could occur would be such events as a natural disaster or a pest outbreak, which could dramatically affect timber supply and thus landowner and forest company decisions regarding production values on their lands.

Benefits of Applied Forward Reasoning

The application of applied forward reasoning has permitted the development of an integrated and interdisciplinary research design in which specific projects can be linked to the broader question of biodiversity conservation. It is indeed ironic that despite forest certification emerging out of a concern for global forest deterioration, there has been very little research or evaluation effort addressing whether the FSC and certification in general are addressing the goals for which they was originally created. The research clearly reveals how a range of analytical, qualitative and quantitative techniques, if carefully chosen and if causal links among mediating factors are conceptualized, can lead to a strong understanding of the impact of forest certification, despite it being a highly complex and dynamic process.

Key futures have been ruled out and only three plausible futures have been identified. This means that future work can focus more narrowly on these scenarios. For instance, a survey of individual consumers would be interesting, but the data reveals that it is unlikely consumers will play any significant role at all in the future of forest certification in the United States.

Our research is still underway. When the floristics and landscape management system work are completed, we will get an even better idea if FSC-style certification leads to any

significant amelioration of biodiversity on lands that were FSC certified. We can presume that if there is either no or marginal differences, then resources might be better spent in other arenas. If, on the other hand, clear impacts are identified, the evidence might encourage FSC supporters to ensure a scenario one style future, in which the FSC has a market presence, and acts as the standard to which others programs are compared.

Discussion of Alternative Analytical Approach

Though we end the presentation of the alternative approach with a discussion of general scenario implications, we wonder about further extensions of the approach. For instance, could some kind of “values” be assigned to the components of the scenarios? Is the approach compatible with econometric modeling?

Taking a step back, we must ask whether we can learn something about the impacts of forest certification we would not have had we not taken this expansive approach. Before the analysis is complete, we can only conjecture. When looking at the analysis from the perspective of making potential recommendations to the FSC, it seems that the approach enables us to better outline strategic concerns for the FSC than would an approach focused only on a limited, static view of the forest certification program. For instance, through the construction of the framework of mediating factors and initial framing of scenarios, we know more about potential indirect effects of the FSC (i.e., via SFI certification) than we would have had we only examined FSC certification itself. Had we chosen to focus on what appear to be the most central factors within the FSC certification system – the FSC standards themselves and the on-the-ground forest impacts from their implementation – we would have missed this potential FSC impact entirely.

In agreement with the philosophy set forth in Bernstein, et al (2000), we intend to be humble about the alternative approach we present and the limitations inherent within it. Here, we talk about some of the major examples of these limitations while inviting others to offer additional ones. Though we have strived to incorporate systematic and rigorous methods within the evaluation, uncertainty still exists around the causal links we identify. This uncertainty may be viewed as increasing at each successive step in the alternative approach. In other words, the uncertainties that are part of the methodologies and analysis of the individual projects will be compounded when we attempt to link findings across these projects.

The scenario plot lines and every other aspect of the approach are limited by our knowledge, our imaginations, and the assumptions we make throughout the analysis. These limitations in addition to time and other resource constraints prevent us from exploring every potentially meaningful combination of mediating factors, predetermined elements, and critical uncertainties. In this way, the scenarios present a simplification of the future, similar to the simplification that is occurring when we map the policy system in the framework of mediating factors. Though our intention is to deal only with general trends, we are, in essence, making some non-quantified point estimates when we create the plot lines so that the different components of the policy system intersect at various points. The scenarios are in no way meant to *predict* the future but rather present alternative futures and help us better understanding the mediating factors and other elements within a policy system (Bernstein, et al, 2000).

This paper has attempted to highlight a gap within environmental policy analysis and illustrate a potential solution. Taking the lead from Bernstein, et al (2000), we have suggested an analytical approach for policy evaluation that will address head-on difficult, real-world policy problems, such as the question of which combination of forest policy approaches will best

support biodiversity conservation in the US. These kinds of problems are highlighted by Parson (2001) and others as pushing the bounds of current analytical capacities.

We suggest that forward reasoning as described by Bernstein and colleagues may be especially applicable to environmental policy problems given the complexity and the data and methodological challenges their analysis entails. An analytical approach that is wide-ranging – especially in the formative stages – and that can address system dynamics is required for the policy basket approaches that have gained increasing prominence within environmental policy-making circles.

Appendix A: Explanatory Work: Scientific Method versus Forward Reasoning

Steps Using the Scientific Method	Steps in Forward Reasoning
1) Propose explanation (independent variables)	1) Identify driving forces
2) Define concepts	2) Specify predetermined elements
3) Formulate hypotheses	3) Identify critical uncertainties
4) Test Hypotheses with experiments <ul style="list-style-type: none"> • Variable group • Control group • Experimental group 	5) Develop scenarios with clear plot lines
5) Refine hypotheses	6) Extract indicators for each scenario
6) Publish, revisit	7) Consider the implication of each scenario
	8) Develop wild cards
Source: Adapted from Bernstein and others [, 2000 #2224]	

Comparison of Key Features	
Scientific Method	Forward reasoning
Empirical	Yes
Subject to empirical verification	Yes, but highly contingent
Non-normative	Evidence, but used for values (i.e. peace)
Transmissible	Only slightly
General	Somewhat
Explanatory	Yes
Provisional	Highly
Source: Adapted from Bernstein and others [, 2000 #2224]	

Appendix B: Applied Policy Analysis versus Applied Forward Reasoning

Steps in Applied Policy Analysis	Steps in Applied Forward Reasoning
3) Problem definition	1) Identify problem
4) State goals	9) Identify goal (s)
10) Identify policy options	11) Identify structure of policy instrument (best to choose one)
4) Consider Impacts of Each Option	12) Develop causal model of key factors that influences whether and how instrument might influence on the ground change
5) Rank Options based on different goals	13) Develop research program addressing different factors
6) Consider implementation and feasibility	14) Identify plausible scenarios and potential impacts on effectiveness
7) Evaluate impacts	15) Identify future research needs
Source: Adapted from Weimer and Vining, Clark	

Comparison of Key Features	
Policy Analysis	Applied Forward reasoning
Empirical	Yes
Contrast different policy instruments	Explore a single policy instrument at a time
Subject to empirical verification	Yes, but highly contingent
Values explicit	Yes, but evidence used for values always contingent and subjective
Transmissible	Transmissible approach, results contingent on basket being explored
Prescriptive	Explanatory and prescriptive
Source: Adapted from Bernstein and others [, 2000 #2224]	

Appendix C: Biodiversity indicators

Goal 1: Biological Diversity Conservation

Beyond the potential threats of chemical pesticides and herbicides and the introduction of exotic species, habitat loss is a primary cause of biodiversity loss. As such, the biological diversity conservation goal of the Doris Duke evaluation concentrates in large part on habitat-related objectives. Due to issues of data unavailability, in the case of the objectives dealing with species of fauna, results will probably be deduced to a great extent from empirical flora results, published reports of other study results in the literature, and other sources that contribute to the researchers' expertise.

Objective x:	To maintain genetic and population diversity of fauna species.
<i>Criterion x:</i>	The implementation of practices beneficial to habitat.
<i>Criterion x:</i>	The presence of viable populations of keystone fauna species.
<i>Relevant Studies:¹</i>	Understory Plant, Overstory Tree, LMS, Precondition/Condition
Objective x:	To protect and encourage the growth of populations of threatened or endangered species of flora and fauna.
<i>Criterion x:</i>	The implementation of practices beneficial to habitat especially critical to threatened or endangered species (e.g., leaving snags, creating wildlife corridors).
<i>Criterion x:</i>	The presence of specialized monitoring procedures for threatened and endangered species.
<i>Relevant Studies:</i>	Understory Plant, Overstory Tree, LMS, Precondition/Condition
Objective x:	To protect environmentally sensitive or high conservation value areas.
<i>Criterion x:</i>	The establishment and implementation of harvest rates and procedures that take into account the special characteristics of the area.
<i>Relevant Studies:</i>	Understory Plant, Overstory Tree, LMS, Precondition/Condition
Objective x:	To promote structural and functional heterogeneity (diversity and redundancy) among floral species.
<i>Criterion x:</i>	The presence of early and late seral habitat.
<i>Relevant Studies:</i>	Understory Plant, Overstory Tree, LMS, Precondition/Condition
Objective x:	To promote the adoption of environmentally friendly methods of pest and noxious plant management.

¹ The key to research component abbreviations is as follows: Yale Structural Ecosystem Feature Study: Overstory Tree Component = "Overstory Tree"; Yale Structural Ecosystem Feature Study: Understory Plant Component = "Understory Plant"; Yale Landscape Management System (LMS) Forecasting Study = "LMS"; Yale Comparative Firm Level Study of Support for FSC = "Firm Level"; Yale Institutional Study Component: Regional Standard-Setting Case Studies = "Regional Standard-Setting"; Yale Institutional Study Component: Standard-Setting Harmonization = "Harmonization"; Yale Institutional Study Component: FSC US and FSC International Coordination = "Federalism & Decision-Making"; Rainforest Alliance Certification Precondition and Condition Comparison Study = "Precondition/Condition"; Bridgespan Organizational Study of FSC = "Organizational"; and Bridgespan Study of Certification Market Conditions = "Market."

Criterion x: A reduction in the usage of chemical pesticides and herbicides.
Relevant Studies: LMS, Precondition/Condition

Objective x: To prevent harmful effects from exotic species.
Criterion x: The presence of monitoring procedures for exotic species.
Relevant Studies: Understory Plant, LMS, Precondition/Condition

Goal 2: Ecological Function Conservation

The ecological function conservation goal focuses mainly on two of the most important services provided by forest ecosystems: clean water and healthy soil resources. The third objective centers on threats to ecological functions through conversion of natural forest to plantation management. The resources and services provided by forest ecosystems depend on certain characteristics, such as spatial distribution of stands and variability in age classes.

Objective x: To prevent soil loss (erosion).
Criterion x: The presence of residual vegetative cover.
Relevant Studies: Understory Plant, Overstory Tree, LMS, Precondition/Condition

Objective x: To safeguard water resources.
Criterion x: A reduction in the amount of sedimentation entering water bodies.
Relevant Studies: Understory Plant, Overstory Tree, LMS, Precondition/Condition

Objective x: To prevent potential detrimental effects from conversion of natural forest to plantations.
Criterion x: Layout of plantation blocks consistent with the patterns of forest stands found within the natural landscape.
Criterion x: Selection of native species when possible.
Relevant Studies: Understory Plant, Overstory Tree, LMS, Precondition/Condition

Bibliography

Endnotes

- Agrawal, A., and C.C. Gibson. 1999. Enchantment and Disenchantment: The Role of Community in Natural Resource Conservation. *World Development* 27 (4):629-649.
- Anderson, James. 1984. *Public Policy Making: An Introduction*. Boston: Houghton Mifflin.
- Anderson, Roy C., and Eric Hansen. 2003. Do Forest Certification Ecolabels Impact Consumer Behavior? Results from an Experiment. Corvallis, OR: Wood Science & Engineering Oregon State University.
- Auld, Graeme, Benjamin Cashore, and Deanna Newsom. 2001. A Look at Forest Certification Through the Eyes of United States Wood and Paper Producers. Paper read at Paper presented to the Auburn Forest Policy Center's conference on Globalization and Private Forestry, March 25-27, at Atlanta, Georgia.
- . 2003. Perspectives on Forest Certification: A Survey Examining Differences Among the US Forest Sectors' Views of Their Forest Certification Alternatives. In *Forest Policy for Private Forestry: Global and Regional Challenges*, edited by L. Teeter, B. Cashore and D. Zhang. Wallingford, UK: CABI Publishing.
- Bernstein, Steven. 2000. Ideas, Social Structure, and the Compromise of Liberal Environmentalism. *European Journal of International Relations* 6 (4 (December)):To be determined.
- . 2002. *The Compromise of Liberal Environmentalism*. New York: Columbia University Press.
- Bernstein, Steven, and Benjamin Cashore. 2000. Globalization, Four Paths of Internationalization and Domestic Policy Change: The Case of Eco-forestry Policy Change in British Columbia, Canada. *Canadian Journal of Political Science* 33 (1):67-99.
- Bernstein, Steven, Richard Ned Lebow, Janice Gross Stein, and Steven Weber. 2000. God Gave Physics the Easy Problems: Adapting Social Science to an Unpredictable World. *European Journal of International Relations* 6 (1):43-76.
- Brewer, Garry. 1974. The Policy Sciences Emerge: To Nurture and Structure a Discipline. *Policy Sciences* 5 (3):239-244.
- Brunner, Ronald. 1991. The Policy Movement as a Policy Problem. *Policy Sciences* 24 (1):65-98.
- Cashore, Benjamin. 2002. Legitimacy and the Privatization of Environmental Governance: How Non State Market-Driven (NSMD) Governance Systems Gain Rule Making Authority. *Governance* 15 (4 (October)):503-529.
- Cashore, Benjamin, Graeme Auld, and Deanna Newsom. 2002. Forest Certification (eco-labeling) and the Impacts of Private Sector Policy Instruments: Comparing Canada, the United States and Europe. Paper read at 16th Biennial Conference of the Association for Canadian Studies in the United States, November 14-18, at San Antonio, Texas.
- . 2003. Forest Certification (Eco-labeling) Programs and their Policy-Making Authority: Explaining Divergence Among North American and European Case Studies. *Forest Policy and Economics* 5 (3).
- . Forthcoming. *Governing Through Markets: Forest Certification and the Emergence of Non-state Authority*. New Haven, CT: Yale University Press.

- Cashore, Benjamin, G. Cornelis van Kooten, Ilan Vertinsky, Graeme Auld, and Julia Affolderbach. 2003. Private or Self-Regulation? A Comparative Study of Forest Certification Choices in Canada, the United States and Germany. *Forest Policy and Economics*.
- Cashore, Benjamin, and James Lawson. 2003. Private Policy Networks and Sustainable Forestry Policy: Comparing Forest Certification Experiences in the US Northeast and the Canadian Maritimes. *Canadian American Public Policy*.
- Cashore, Benjamin, and Ilan Vertinsky. 2000. Policy Networks and Firm Behaviours: Governance Systems and Firm Responses to External Demands for Sustainable Forest Management. *Policy Sciences* 33 (March):1-30.
- Cashore, Benjamin, Ilan Vertinsky, and Rachana Raizada. 2001. Firm Responses to External Pressures for Sustainable Forest Management in British Columbia and the US Pacific Northwest. In *Sustaining the Pacific Coast Forests: Forging Truces in the War in the Woods*, edited by D. Salazar and D. Alper. Vancouver: UBC Press.
- Clark, Tim. 2002. *Natural Resources Management Analysis for the Natural Resources Professional*. New Haven: Yale University Press.
- Clark, Tim W. 1997. *Averting extinction : reconstructing endangered species recovery*. New Haven, Conn. ; London: Yale University Press.
- . 2001. *Species and Ecosystem Conservation: An Interdisciplinary Approach*: Yale University Press.
- Clark, Tim, Andrew Williard, and Christina Cromley, eds. 2000. *Foundations of Natural Resources Policy and Management*. New Haven: Yale University Press.
- Collier, David. 1991. The Comparative Method: Two Decades of Change. In *Comparative Political Dynamics*, edited by D. A. R. a. K. P. Erickson. New York: Harper Collins.
- Conroy, Michael. 2001. Can Advocacy led Certification Systems transform Global Corporate Practices? Evidence and Some Theory.
- Cutler, Claire, Virginia Haufler, and Tony Porter, eds. 1999. *Private Authority in International Politics*. New York: SUNY Press.
- DeLeon, Peter. 1999. The Stages Approach to the Policy Process: What Has It Done? Where Is It Going? In *Theories of the Policy Process*, edited by P. Sabatier. Boulder: Westview Press.
- deLeon, Peter. 1999. The Stages Approach to the Policy Process: What Has it Done? Where is it Going? In *Theories of the Policy Process*, edited by P. Sabatier. Boulder, CO: Westview Press.
- Elliott, Chris, and Rodolphe Schlaepfer. 2001. The Advocacy Coalition Framework: Application to the Policy Process for the Development of Forest Certification in Sweden. *Journal of European Public Policy* 8 (4):642-661.
- Environmental Protection Agency. United States. 1997. The Benefits and Costs of the Clean Air Act, 1970 to 1990.
- . 2003. FY 2002 Annual Report.
- Esping-Andersen, Gosta. 1989. The Three Political Economies of the Welfare State. *Canadian Review of Sociology and Anthropology* 26 (1):10-36.
- . 1990. *Three Worlds of Welfare Capitalism*. Princeton, NJ: Princeton University Press.
- Esty, Daniel C. 2002. Why Measurement Matters. In *Environmental Performance Measurement: The Global Report 2002-2003*, edited by D. Esty and P. Cornelius. New York: Oxford University Press.

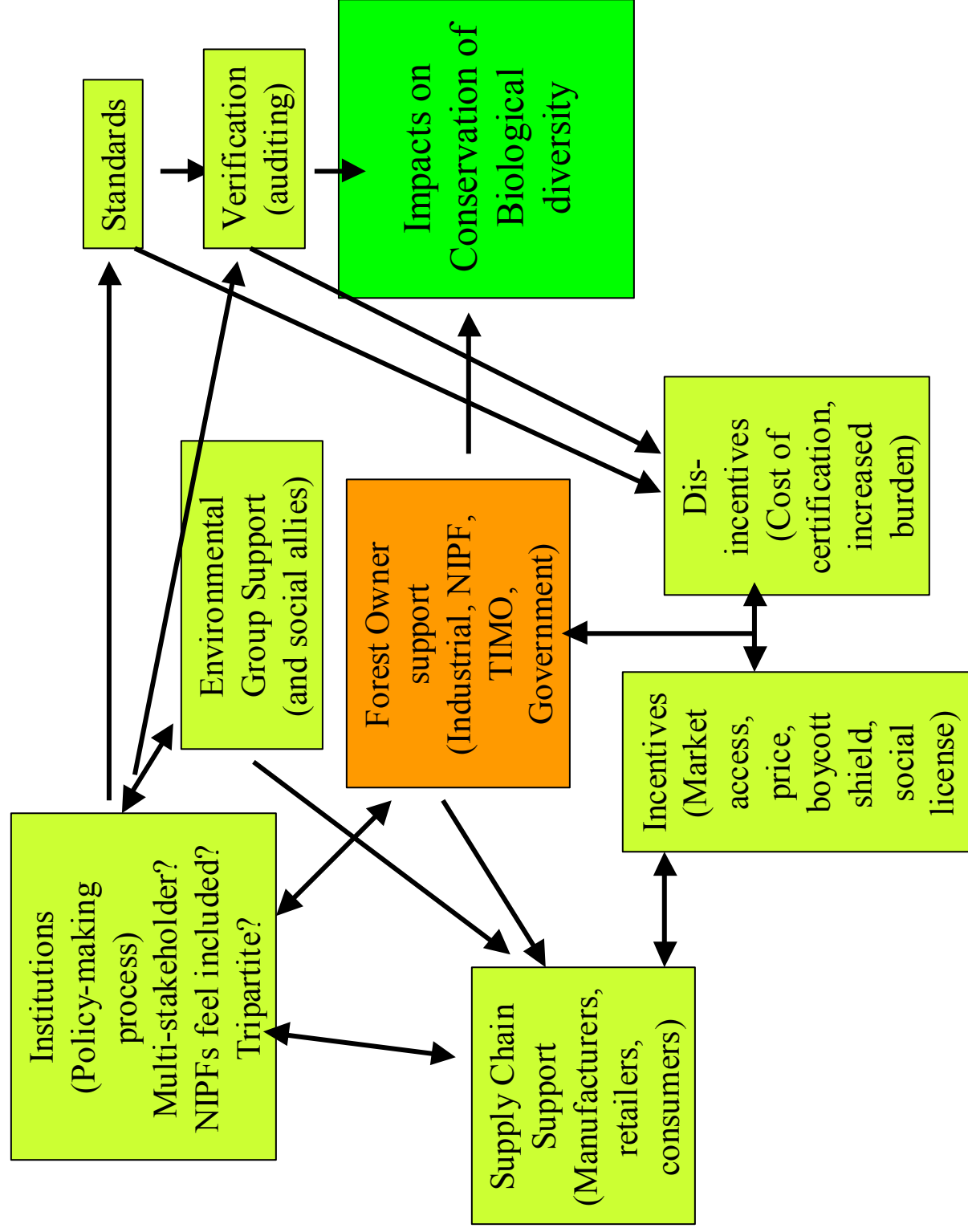
- Esty, Daniel C., and Peter Cornelius, eds. 2002. *Environmental Performance Measurement: The Global Report 2002-2003*. New York: Oxford University Press.
- Florini, Ann. 2003. *The coming democracy : new rules for running a new world*. Washington, DC: Island Press.
- Gereffi, Gary, Ronie Garcia-Johnson, and Erika Sasser. 2001. The NGO-Industrial Complex. *Foreign Policy* (July/August):56-65.
- Gunningham, Neil, Peter N. Grabosky, and Darren Sinclair. 1998. *Smart regulation - designing environmental policy, Oxford socio-legal studies*. Oxford and New York: Clarendon Press and Oxford University Press.
- Gunningham, Neil, and Darren Sinclair. 2002. *Leaders and Laggards: Next-Generation Environmental Regulation*. Australia: Greenleaf Publishing.
- Hacker, Jakob. 2001. *The Divided Welfare State*. New York: Cambridge University Press.
- Harrison, Kathryn. 1996. The Regulator's Dilemma: Regulation of Pulp Mill Effluents in the Canadian Federal State. *Canadian Journal of Political Science* 29 (3 (September)):469-497.
- . 1999. Racing to the Top or the Bottom? Industry Resistance to Eco-labelling of Paper Products in Three Jurisdictions. *Environmental Politics* 8 (4).
- Haufler, Virginia. 2001. *A Public Role for the Private Sector: Industry Self-Regulation in a Global Economy*. Washington, DC: Carnegie Endowment for International Peace.
- Howlett, Michael. 2000a. Managing the "Hollow State": Procedural Policy Instruments and Modern Governance. *Canadian Public Administration* 43 (4):412-431.
- . 2000b. Managing the 'hollow state': procedural policy instruments and modern governance. *Canadian Public Administration* 43 (4):412-432.
- . 2002. Do Networks Matter? Linking Policy Network Structure to Policy Outcomes: Evidence from Four Canadian Policy Sectors. *Canadian Journal of Political Science* 35 (2 (September)):235-268.
- Howlett, Michael, and M. Ramesh. 1995. *Studying Public Policy Cycles and Policy Subsystems*. 1 vols. Toronto: Oxford University Press.
- Jennings, P. Devereaux, and Paul A. Zandbergen. 1995. Ecologically Sustainable Organizations: An Institutional Approach. *Academy of Management Review* 20 (4):1015-1052.
- Kernaghan. 1999. Voluntary Approaches, the Environment and the Law: A Canadian Perspective. In *Voluntary Approaches in Environmental Policy*, edited by C. C. a. F. Leveque: Kluwer.
- Kerr, Greiner, Anderson, and Inc. April. 1999. Analysis and Evaluation of the EPA Common Sense Initiative: Prepared for US EPA.
- Knapp, G., and T. Kim, eds. 1998. *Environmental Program Evaluation: A Primer*. Urbana: University of Illinois Press.
- Lawson, James, and Benjamin Cashore. 2001. Firm Choices on Sustainable Forestry Forest Certification: The Case of JD Irving, Ltd. In *Forest Policy for Private Forestry*, edited by B. C. L. Teeter, and D. Zhang. Wallingford, UK: CABI Publishing.
- Marsh, David, and R.A.W. Rhodes. 1992a. Policy Communities and Issue Networks: Beyond Typology. In *Policy Networks in British Government*. Oxford: Clarendon Press.
- , eds. 1992b. *Policy Networks in British Government*. Oxford: Clarendon Press.
- Meidinger, Errol E. 1999. "Private" Environmental Regulation, Human Rights, and Community. *Buffalo Environmental Law Journal* 7:123-237.

- . 2001. Environmental Law for Global Civil Society: the Forest Certification Prototype. Paper read at Paper prepared for the Conference on Social and Political Dimensions of Forest Certification, June 20-22, at Freiburg, Germany.
- Metrick, Andrew, and Martin Weitzman. 1998. Conflicts and Choice in Biodiversity Protection. *The Journal of Economic Perspectives* 12 (3):21-34.
- National Academy of Public Administration. 2000. Transforming Environmental Protection for the 21st Century. Washington: National Academy of Public Administration.
- Newsom, Deanna, Benjamin Cashore, Graeme Auld, and Jim Granskog. 2003. Forest Certification in the Heart of Dixie: A Survey Of Alabama Landowners. In *Forest Policy for Private Forestry: Global and Regional Challenges*, edited by L. Teeter, B. Cashore and D. Zhang. Wallingford, UK: CABI Publishing.
- North, Douglas. 1981. *Structure and Change in Economic History*. New York, NY: Norton and Company.
- Oliver, Chadwick. 2003. What is Sustainable Forestry, How do we Achieve it? *Journal of Forestry*.
- Ostrom, E. 1998. Institutional Rational Choice: An Assessment of the Institutional Analysis and Development Framework. In *Theories of the Policy Process: Theoretical Lenses on Public Policy*, edited by P. Sabatier. Boulder, Colorado: Westview press.
- Pal, Leslie. 1997. *Beyond Policy Analysis : Public Issue Management in Turbulent Times*. Scarborough, Ont.: ITP Nelson.
- Parson, Ted. 2002. Introduction. In *Governing the Environment*. Toronto: University of Toronto Press.
- Patton, Michael. 2002. *Qualitative Research and Evaluation Methods*. Third ed. Thousand Oaks: Sage Publications, Inc.
- Pierson, Paul. 1993. When Effect Becomes Cause: Policy Feedback and Political Change. *World Politics* 45 (4 (July)):595-628.
- . 2000. Increasing Returns, Path Dependence, and the Study of Politics. *American Political Science Review* 94 (2):251-268.
- Prakash, Aseem. 1999. A New-Institutional Perspective on ISO 14000 and Responsible Care. *Business Strategy and the Environment* 8:322-35.
- . 2000. *Greening of the firm: The politics of corporate environmentalism*. Cambridge: Cambridge University Press.
- Rivera, Jorge. 2002. Assessing a voluntary environmental initiative in the developing world: The Costa Rican Certification for Sustainable Tourism. *Policy Sciences* 35 (4).
- . 2003. Institutional pressures and voluntary beyond-compliance environmental behavior in developing countries: Evidence from Costa Rica. *Under Review at Society and Natural Resources*.
- Rogers, Patricia, Timothy Hacsí, Anthony Petrosino, and Tracy Huebner, ed. 2000. *Program Theory in Evaluation: Challenges and Opportunities*. Vol. 87, *New Directions for Evaluation*. San Francisco: Jossey-Bass.
- Sabatier, Paul. 1991. Toward Better Theories of the Policy Process. *PS* 29.
- , ed. 1999a. *An Advocacy Coalition Lens on Environmental Policy*. Cambridge, Mass: MIT Press.
- , ed. 1999b. *Theories of the Policy Process*. Boulder, CO: Westview Press.
- Salaman, Lester M. 2002. Introduction. In *The Tools of Government: A Guide to the New Governance*. New York: Oxford University Press.



- Salamon, Lester. 1981. Rethinking Public Management: Third-Party Government and the Changing Forms of Government Action. *Public Policy* 29 (3):255-275.
- Sasser, Erika N. 2002. The Certification Solution: NGO Promotion of Private, Voluntary Self-Regulation. Paper read at 74th Annual Meeting of the Canadian Political Science Association, May 29-31, 2002, at Toronto, Ontario.
- . 2003. Gaining Leverage: NGO Influence on Certification Institutions in the Forest Products Sector. In *Forest Policy for Private Forestry*, edited by L. Teeter, B. Cashore and D. Zhang. Oxon, UK: CAB International.
- Schön, D., M. Rein. 1994. *Frame Reflection*. New York: BasicBooks.
- Skocpol, Theda. 1986. Bringing the State Back In: Strategies of Analysis in Current Research. In *Bringing the State Back In*, edited by P. B. Evans, D. Rueschemeyer and T. Skocpol. Cambridge: Cambridge University Press.
- . 1995. Why I am an Historical Institutionalists. *Polity* 28 (1):103- 106.
- The Heinz Center. 2002. *The State of the Nation's Ecosystems: Measuring the Lands, Waters, and Living Resources of the United States*. Cambridge: University of Cambridge.
- Tribe, Laurence H. 1982. Policy Science: Analysis or Ideology. *Philosophy and Public Affairs* 2:66-110.
- Van Waarden, Frans. 1992. Dimensions and Types of Policy Networks. *European Journal of Political Research* (21):29-52.
- Vertinsky, Ilan B., and Charlene Zietsma. 1998. Corporate Greening and Environmental Protection Performance: Static and Dynamic Analysis. Vancouver: Faculty of Commerce and Business Administration.
- Weaver, R. Kent, and Bert A. Rockman, eds. 1993. *Do Institutions Matter?* 1 vols. Washington, D.C.: The Brookings Institution.
- Weimer, David L., and Aidan R. Vining. 1999. *Policy Analysis: Concepts and Practice, 2d edition*. Englewood Cliffs, NJ: Prentice-Hall.
- Weitzman, Martin. 1998. Prices or Quantities: Policy Instrument Choice Under Uncertainty. *The Journal of Economic Perspectives*.
- Young, Oran. 1994. *International Governance: Protecting the Environment in a Stateless Society*. Ithica, New York: Cornell University Press.
- Young, Oran R. 1999. *The effectiveness of international environmental regimes : causal connections and behavioral mechanisms*. Cambridge, Mass.: MIT Press.
- Zietsma, Charlene, and Ilan B. Vertinsky. 1999-2001. Shades of Green: Cognitive Framing and the Dynamics of Corporate Environmental Response. *Journal of Business Administration and Policy Analysis* 27-29:217-247.

Appendix D: Causal Model of mediating factors that influence impact of Forest Certification on Biodiversity Conservation

12/18/2003 Draft - Do Not Cite Without Permission



ATTACHMENT A
Research Component Contributions to Describing and Analysing

	 Primary contribution	 Potential contribution	<i>Biological Diversity Conservation</i>	<i>Ecological Function Conservation</i>	<i>Forest Owner Support for FSC</i>	<i>Institutional Governance</i>	<i>Supply Chain Support</i>	<i>Economic Incentives and Disincentives</i>	<i>Standards Content</i>	<i>Auditing</i>
LMS Study										
Understory Plant Study										
Overstory Tree Study										
Precondition/Condition Study										
Market Study										
Organizational Study										
Firm Level Study										
Regional Standard-Setting Study										
Harmonization Study										
Federalism and Decision-Making Study										
Research on Consumer Attitudes (not begun yet)										
Survey of Alabama Land Owners										
Survey of US Forest Sector										
Comparative Research on Different Countries										

Endnotes

¹ Of Young's (Young 1994: 142-152) six measures of effectiveness, only one addressed whether the regime actually solved the problem for which it was created. The six definitions he offers are: *Success in Problem Solving*; *Degree of Goal Attainment*; *Effecting Behavior Change*; *Process Success*; *Constitutive Effectiveness*; *Evaluative Effectiveness*. For an updated treatment see Young (1999). O'Neill, et al. (2003) add another definition of effectiveness, arguing that cooperation inherent in multi-lateral environmental agreement (MEA) processes should be assessed for value beyond that of the MEA outputs can be measured (e.g., program compliance rates). These authors advocate a move beyond a functional definition of effectiveness toward a wider understanding of the impacts of multinational cooperation on transmitting norms and policy ideas and otherwise affecting broadly defined social change.

² Certainly governments must do this all the time, but usually they do so using only a limited number of reference points, and rarely identify systematically the range of factors that might mediate pressure in the long run

³ Brewer was a student of Harold Lasswell, who is largely attributed with creating the discipline of policy sciences. Policy science was distinct in that it was meant to be multi-disciplinary, problem-solving, and admittedly normative (Clark, et al, 2000). As Howlett and Ramesh explain, Lasswell and colleagues conceived of policy science as being oriented toward real-world problems and not engaging in "purely academic" debates (Howlett and Ramesh, 1995: 3). Lasswell states that the "policy scientist is far more interested in evaluating and reconstructing the practice of society than in . . . the higher abstractions from which the values are derived" (1951: 3). As Howlett and Ramesh indicated, Lasswell and the other founders intended that policy science should not be cloaked in the guise of 'scientific objectivity' but should recognize the impossibility of separating goals and means, or values and techniques, in the study of government actions" (Howlett and Ramesh, 1995: 3). Another student of Lasswell's downplayed a purely scientific popperian approach: the "purpose of the policy sciences as 'science' is to realize more of the potential for free choice through the sharing of insight . . . The purpose is not prediction" (emphasis in original) (Brunner, 1991). The approach has gone through adjustment with Garry Brewer (1974), revising the initial list of steps. There are also distinctions and overlaps between the actual activity of policy-making and that of analyzing it (Browne and Wildasvsky, 1984; DeLeon, 1999).

⁴ There are important exceptions. Kent Weaver's work on how to best design electoral systems does assess potential impacts, though here impacts are defined in ways political scientists can measure (such as equity of representation and articulation of issue), rather than on whether they would reverse deterioration of natural ecosystems. In addition, Eleanor Ostrom's work with colleagues most certainly does address institutional designs, including how they might impact deforestation in developing countries [cite forests book]. However, even in this case institutional designs are addressed for their ability to influence individual behavior, with a strong emphasis on easy to measure indicators such as forest cover and amount of harvesting. Less attention is paid to understanding the complex and biologically diverse functioning ecosystems and if, for example, whether an emphasis on reducing extraction might downplay other important ecosystem structures and functions.

⁵ For an important exception see Wargo (1998).

⁶ Because evaluation as a separate discipline focuses most often on specific cases of implementation of larger policies (i.e., programs), it is frequently known as program evaluation. We use this term within the paper to avoid confusion between the activities we allocate to evaluation and those within more general policy analysis.

⁷ These techniques apply whether the evaluation seeks to assess the performance of a program against its intended goals (i.e., an outcome or impact evaluation) or whether it has been implemented according to its stated guidelines (i.e., a process evaluation) (Howlett and Ramesh, 1995). A process evaluation can also be thought of as one that investigates the factors behind the performance a program achieves (Knaap and Kim, 1998).

⁸ The Heinz Center. 1999. Designing a Report on the State of the Nation's Ecosystems. (Washington: H. John Heinz III Center). Available at www.us-ecosystems.org.

⁹ It reports on ten major kinds indicators for six major ecosystem types (coasts and oceans; farmlands; forests; fresh waters; grassland and shrublands; and urban and suburban areas). The ten kinds of indicators cover the physical dimensions of the systems, their chemical and physical conditions, the status of their biological components, and the amounts of goods and services people receive from them. It acknowledges that adequate data to report on the indicators is not always available. Where possible, the researchers provide context for how indicators have changed over time (i.e., trends), how indicators differ by region, and how the data compare to widely accepted reference points (i.e., regulatory standards for chemical contaminants).

¹⁰ See: The individual studies were analyzed together to produce the NAPA report entitled, *environment.gov: Transforming Environmental Protection for the 21st Century*. This report followed two others that had also been commissioned by Congress on the subject of policy innovations (*Setting Priorities, Getting Results: A New Direction for EPA* (1995) and *Resolving the Paradox of Environmental Protection: An Agenda for Congress, EPA, and the States* (1997)). In *environment.gov*, NAPA summarizes findings from the individual studies and uses the findings to recommend approaches for broader implementation.

A study entitled, “Green Permits and Cooperative Agreements: A Report on Regulatory Innovation Programs in Oregon and Washington” (Speir, 2000), used a comparative analytical approach to examine two state environmental policy programs. Case studies were performed on a few different participants within each program. Findings and recommendations focused on the accountability, efficiency, and flexibility shown by the program in the case studies, in accordance with these goals stated by the program founders. In examining the case study experiences, Speir considered the influence of stakeholders and of state and federal government relations on the dynamics within the programs. The overall evaluative approach can be considered one of an effectiveness or impact assessment. Because the outcome goals of accountability, efficiency, and flexibility are also process characteristics in and of themselves, the study has elements of a process evaluation as well. The analysis did not consider the environmental management standards or environmental outcomes related to the green permits and cooperative agreements being described. The newness of both programs is one reason why a more comprehensive evaluation may not have been possible.

Air Pollution Trading Programs

The purpose of “Analysis of Volatile Organic Compound Air Pollution Trading Systems” was to examine the extent to which five different programs achieved their goals with regard to environmental and economic outcomes (Hix, et al, 2000). The overall approach is that of a process as well as effectiveness, or impacts, evaluation with some attention paid to factors outside of the boundaries of the program. Through interviews and secondary data review, the researchers addressed a series of questions about the design, implementation, and effects of the programs under consideration. When analyzing the activity under the programs, they compare them to the federal sulfur dioxide trading program, which many consider to be the most successful air pollution trading program. The study takes outside factors into account by looking at those forces that persuade or dissuade companies from making trades.

Because of a lack of regular monitoring data on the programs, Hix and colleagues calculate the economic costs and benefits of the programs “qualitatively” (Hix, et al, 2000). They do not quantify these costs or benefits but rather make general observations on them stemming from data they collected in their interviews and secondary data collection efforts. With regard to the environmental benefits, the researchers are able to report emissions reductions corresponding to offset ratios and expiration of pollution credits. They do not take the data to the next level by considering the reduced illness or other outcomes associated with these emission reductions. An important feature of the study is its qualitative analysis of the emissions reductions in comparison to what would have been achieved under the status quo public policy approaches.

ISO 14001 and Regional Innovation Program

The third example entitled, “ISO 14001 and EPA’s Region 1’ StarTrack Program,” (Nash, et al, 2000) is a study of those two programs. The overall approach is one of a process and effectiveness, or impact, analysis for each program, with some attention paid to factors outside of the programs. For ISO 14001, the researchers prepare a series of case studies and also perform a statistical analysis of the characteristics of ISO 14001-certified companies versus those that are not. Nash and colleagues were particularly interested in 1) why certain companies chose to participate in ISO 14001 and others did not and 2) if ISO 14001 participation created any differences in environmental performance. The Toxics Release Inventory (TRI) and a few other databases reporting chemical releases along with the information provided by company representatives helped the researchers in answering the two questions. They do not attempt to link performance differences to differences in ultimate environmental outcomes (e.g., reduced illness associated with air pollution)

The StarTrack component of the analysis is performed using all qualitative data. The researchers interviewed the majority of the program’s participants as well as other stakeholders (e.g., agency staff, community members). Again, they focused on why companies had decided to take part and what environmental management changes had taken place as a result. The study does not attempt to present environmental practice changes or results in the aggregate for the companies. While the reason for looking at these two programs in particular within the same study was not made very clear, the researchers do discuss how these programs might interact with one another based on their findings (e.g., being part of StarTrack encouraged companies to go for ISO 14001).

¹¹ Its approach integrated effectiveness to process and impact. It relied on qualitative data, primary data collection (e.g., interviews) and secondary sources (e.g., meeting minutes).

Researchers characterized the outputs of the project in terms of aggregate pilot projects initiated, alternative regulation suggested, etc. Much of the attention was given to process elements that were at the core of the CSI goals to enable an environment of collaborative, “out-of-the-box” environmental policy-making. The researchers assessed the extent to which this happened across all of the projects and the key factors helping and hindering it. A good deal of emphasis was also placed within the approach on the dynamic nature of the program. The researchers traced how the program had evolved over the years and considered the future of the program in relation to factors – both in side and outside – CSI that they considered important.

¹² For an exception, see Anderson and Hansen (2003).

¹³ While auditing for these programs does look at a variety of ecological issues, such as riparian zone management, wildlife policy, and protected areas, the results are reported as a binary classification of certified or non-certified (Meidinger, 2001).

¹⁴ NOTE: Our research efforts began before this evaluation request and continue independent of it as well. For clarity, we include all of our research efforts below, regardless of from where funding came. The research began at Auburn University with strong support from Dean Brinker, and Auburn's Center for Forest Sustainability and Forest Policy Center. Two Masters Students, Deanna and Newsom and Graeme Auld, were recruited there who continue to do collaborative work with me. Also hired Jamie Lawson, now Professor at Memorial University. Other university partners for ongoing and future projects include Errol Meidinger SUNY Buffalo; Fred Gale, University of Tasmania, Australia; Aseem Prakash, University of Washington; Erika Sasser, Duke University. Current faculty research collaborations include Chad Oliver and Mark Ashton, and Rob Mendelsohn. Future work is anticipated with recent faculty additions e Lloyd Irland and Michael Conroy. Other faculty interest in tropics and certification include Lisa Curran and Florencia Montagnini. We have significant research staff and student support including Emily Noah, Megan Mattox, Alex Finkral, Dave Ellum, Graeme Auld, Liz Gordon, and Barb Bamberger.

¹⁵ The goals put forth are biodiversity conservation, ecological function conservation, economic incentives, support for FSC, institutional responsiveness, and efficiency. Within each goal, we develop more specific objectives and criteria. For example, under the goal of institutional responsiveness, we will measure the extent to which the FSC achieves the objective of ensuring consultation with all relevant stakeholders. The corresponding criterion is a check for the presence of processes that enable relevant stakeholders to join or interact with the certification system.

¹⁶ For the understory analysis, measures calculated include species richness and composition. The overstory analysis focuses on leaf area and growth efficiency.

¹⁷ We define prescriptiveness as an index of three different components contained within each standard: level of detail, discretion, and degree of change. The prescriptiveness analysis is to be performed using two different instruments, a prescriptiveness categorization matrix based on the three components and a comparative prescriptiveness decision tree that enables direct comparisons between two different sets of standards.

¹⁸ For example, the extent of landowner participation in the different regional standard-setting processes will be investigated for potential consequences on the feasibility of the forest practices indicators as they are implemented in the different US regions.

¹⁹ The idea of an effect whereby FSC certification indirectly causes SFI standards to be "raised" is explored within Cashore, et al, Forthcoming.