

Evaluation of the impacts of Forest Stewardship Council (FSC) certification of natural forest management in the tropics: a rigorous approach to assessment of a complex conservation intervention

C. ROMERO^{a,c}, E.O. SILLS^{b,c}, M.R. GUARIGUATA^c, P.O. CERUTTI^c, G. LESCUYER^c and F.E. PUTZ^{a,c}

^aDepartment of Biology, University of Florida, 220 Bartram Hall, Gainesville, FL 32611-8526, USA

^bDepartment of Forestry and Environmental Resources, North Carolina State University, Durham, NC, USA

^cCenter for International Forestry Research (CIFOR), Jl. CIFOR, Situ Gede, Sindang Barang, Bogor 16680, Indonesia

Email: romero@ufl.edu, sils@ncsu.edu, m.guariguata@cgiar.org, p.cerutti@cgiar.org, g.lescuyer@cgiar.org, fep@ufl.edu

SUMMARY

After more than 20 years and substantial investments of time and money, Forest Stewardship Council (FSC) certification of tropical forest management is due for a stringent impact evaluation. For any social, ecological, and economic outcomes to be attributed to FSC certification, rival explanations need to be ruled out. We recognize that different types of knowledge about FSC impacts derived from information gathered through a range of methods can satisfy the evidence-needs of different stakeholders. But this paper describes a roadmap based on rigorous methods to assess whether FSC certification delivers on its expected outcomes and the underlying mechanisms through which changes can be attributable to FSC. To this end, background studies that provide contextual knowledge related to implementation of FSC certification are proposed to account for any positive self-selection biases and to capture the temporal dynamics of certification including changes in the sociopolitical and economic contexts that influence certification decisions.

Keywords: impact evaluation, conservation interventions, tropical forestry, market-based conservation

Evaluation des impacts de la certification Forest Stewardship Council (FSC) d'une gestion de forêts naturelles dans les tropiques: une approche rigoureuse d'estimation d'une intervention complexe de conservation

C. ROMERO, E.O. SILLS, M.R. GUARIGUATA, P.O. CERUTTI, G. LESCUYER et F.E. PUTZ

Après plus de 20 ans et des investissements substantiels de temps et d'argent, une évaluation d'impact rigoureuse doit être conduite du schéma de certification Forest Stewardship Council (FSC) de la gestion des forêts tropicales. Pour que des impacts sociaux, écologiques et économiques soient attribués à la certification FSC et à elle seule, les explications concurrentes doivent être écartées. Nous reconnaissons que différents types de connaissances sur les impacts du FSC, tirés des résultats obtenus grâce à une gamme de méthodes différentes, peuvent satisfaire les besoins en données probantes de certaines parties prenantes. Cependant, cet article décrit une feuille de route basée sur des méthodes rigoureuses pour évaluer si la certification FSC produit les impacts escomptés et les mécanismes sous-jacents par lesquels les changements peuvent être attribués au FSC. À cette fin, des études qui fournissent des connaissances contextuelles liées à la mise en œuvre de la certification FSC sont proposées pour tenir compte de tout biais positif d'auto-sélection (*self-selection*) et pour saisir la dynamique temporelle de la certification, y compris les changements dans les contextes sociopolitiques et économiques qui influencent les décisions de certification.

Evaluación de los impactos de la certificación de manejo forestal de bosques naturales tropicales por el Forest Stewardship Council (FSC): aproximación rigurosa para la estimación de una intervención de conservación compleja

C. ROMERO, E.O. SILLS, M.R. GUARIGUATA, P.O. CERUTTI, G. LESCUYER y F.E. PUTZ

La certificación del manejo de bosques naturales a través del Forest Stewardship Council (FSC) se ha puesto en marcha desde hace más de 20 años. A pesar de las sustanciales inversiones de tiempo y dinero en este mecanismo, aún no se cuenta con una evaluación rigurosa de los impactos de su adopción. La atribución de cualquier impacto social, ecológico y económico a la certificación FSC requiere descartar el efecto que otros factores puedan tener en explicar los resultados observados. Reconocemos que los diferentes actores sociales que participan en la

certificación forestal tienen diversas necesidades de conocimiento sobre los impactos del FSC. Aún así, este documento describe una hoja de ruta basada en métodos rigurosos para evaluar si la certificación FSC cumple con los resultados esperados y para elucidar los mecanismos subyacentes a través de los cuales los cambios observados puedan atribuirse al FSC. Con este fin, se delinea el contenido de una serie de estudios los cuales incluyen análisis de antecedentes que aportan conocimientos contextuales relacionados con la adopción de la certificación FSC tales como cambios en los contextos sociopolíticos y económicos que influyen en las decisiones de la gestión de bosques y su certificación. Estos estudios contribuyen a entender la existencia de sesgos de auto-selección positiva en la participación en el esquema FSC, así como a describir la dinámica temporal de la certificación.

INTRODUCTION

Certification of responsible (i.e. environmentally sound, socially equitable, financially viable, and politically transparent) forest management by the Forest Stewardship Council (FSC, www.fsc.org), which has been underway since 1993, is promoted as a way to maintain and enhance forest values. It encompasses a multi-layered intervention applied in complex and changing contexts that involve a range of interested parties all along the product chain. It includes consideration of issues and processes from rights to resources to purchases of manufactured products. Over the 20-plus years since the FSC was founded, the deleterious environmental, socio-economic, and political impacts of conventional agriculture and forestry have become more apparent. As a path-breaking approach to conservation, FSC certification is used as a template for other interventions designed to promote responsible management of renewable natural resources (Chaplin-Kramer *et al.* 2015).

FSC certification promotes forest conservation through sound use. Given limited budgets for promoting responsible forest management and the many prominent failures in this endeavor (e.g. the Tropical Forest Action Plan – TFAP, Colchester and Lohman 1990, Winterbottom 1990, Pfaff *et al.* 2013), it is critical to identify the most effective programs and initiatives (hereafter, interventions) and the conditions under which their effectiveness was realized. It is also clear that there is no ‘silver bullet’ for reconciling conservation and development: some forest policy interventions might work well when implemented in a particular way to achieve particular outcomes at a particular time and place, but fail miserably under other circumstances.

The continued high rates of forest degradation due to poor forestry practices motivate efforts to determine how, where, and under what circumstances FSC certification has contributed to the maintenance and improvement of forest values (Visseren-Hamakers and Pattberg 2013). Our interest, and the focus of this paper, is on rigorous, independent methods employed to evaluate FSC’s delivery of its expected outcomes and on the mechanisms through which change was possible, when and why change occurred, who benefited, and with what cost-benefit ratios (Vincent 2016). The various available approaches to assessment of programs like certification vary in their costs as well as robustness, but all can be suitable for a range of purposes that vary with the needs of stakeholders (Bamberger *et al.* 2012). For example, companies that source FSC timber might be satisfied knowing that the timber comes

from properly managed forest, irrespective of whether responsible management resulted from adoption of FSC certification. For this particular type of stakeholder, robust counterfactual analyses (i.e. what would have been the outcome of a certified unit had it not received certification; Ferraro 2009) might not be warranted. We recognize that in this case companies might not mind spending funds to remain certified, whether or not this decision translates into any demonstrable sustainability, social welfare, or other benefits. Similarly, some stakeholders might be satisfied by reports based on interviews with people very much involved in FSC certification insofar as those interviews reveal insights about the shortcomings of the intervention that can then be improved through corrective measures (e.g. Van Kuijk *et al.* 2009). Although these sorts of stakeholders might consider full-blown impact evaluations excessively expensive and unnecessary, we believe it important to remember that certification is a private, voluntary, market-driven governance-focused intervention that ultimately aims to improve forest management. We believe that it is essential to understand whether this intervention is having the expected impacts and how it might be improved.

In general terms, evaluations should consider impacts that are positive and negative, direct and indirect, short- and long-term, and intended and unintended (OECD 2002). These impacts will surely vary among regions and participants as well as over time due in part to the heterogeneity in implementation of the intervention as well as to changes in critical contextual factors including legal frameworks and social processes. Well-structured, verifiable, independent, and objective information about both implementation and impacts can inform negotiations, decision-making, and resource allocations to future interventions that aim to promote socially responsible, economically viable, ecologically sound, and politically accountable resource management.

The remainder of this article describes a roadmap for evaluation of the impacts of FSC certification of managed natural tropical forests. Certification impacts are changes in the forest management unit (FMUs: the entities that get certified such as concessions, communities or private lands), neighboring communities, forest workers, and local and national stakeholders. We emphasize that determining whether and which outcomes can be attribute to FSC certification requires detailed understanding of context and process, and thus we call for a multi-pronged research effort with broad stakeholder input. Many elements of the proposed roadmap could be adapted to guide the evaluation of other sustainability-promoting interventions, particularly those that rely on

voluntary adoption.¹ The recommended approach might also be used to assess more traditional, state-driven policy interventions (e.g. the adoption and implementation of forest management plans in logging concessions) in order to better understand and improve their impacts.

ON EVALUATION

Basic traits of sound evaluations

Impact evaluations are becoming a priority for donors, governments, non-governmental organizations, academics, and conservation practitioners who work towards socio-environmental sustainability (Ferraro and Pattanayak 2006). Despite a growing body of studies that show what works for conservation (e.g. Ferraro and Pressey 2015), much that remains to be learned about evaluation of conservation interventions can be garnered from the more substantial experience with evaluation of development, health, and education interventions (e.g. Campbell Collaboration).

It is often claimed that impact evaluation and other sorts of evidence-based learning can improve policies, especially if informed by analyses of the drivers of policy change (Eklin *et al.* 2014, Resnick 2015). Indeed, evaluation is recognized as a key step in the policy process (Bell *et al.* 2011). Understanding how policies are made and applied in regards to a specific case-study, such as FSC certification, is central to reflections on whether and how research can influence policy (Nason *et al.* 2007). This understanding is, in turn, important to promote synergies among interventions such as between FSC and REDD+ (Lambin *et al.* 2014) or among different modes of governance (Heilmayr and Lambin 2016). Researchers and practitioners have stressed the need to formalize lessons learned so as to more efficiently use the scarce resources available for conservation (Ferraro and Pattanayak 2006). The combined communities of conservationists and evaluators are striving to determine how to conduct sound and robust evaluations so as to make evident the full range of successes and failures (Baylis *et al.* 2015, Ezzine de Blas *et al.* 2016, Le Velly and Dutilly 2016, Mascia *et al.* 2014, McKinnon *et al.* 2015a, 2015b, Milder *et al.* 2015, Miteva *et al.* 2012, SCR 2012).

Effective evaluations are carried out in transparent manners, with integrity and inclusiveness so that results address accountability concerns (Farley *et al.* 2012, Rogers 2012). Such evaluations can generate knowledge for a range of people and institutions interested in and affected by implementation of the intervention (Romero and Castrén 2013, Romero *et al.* 2013). This logic is based on recognition of the evaluation endeavor as an opportunity to learn, a vehicle for inclusion of various stakeholders, a space for reflection and

deliberation, and a mechanism to reveal and then learn from past mistakes and successes.

Given the costs of evaluations and the high stakes, evaluators should be independent and unbiased researchers without preferences about the outcomes of the assessment (GAO 2009, Gertler *et al.* 2011, Perrin 2012, PROFOR and FAO 2011, Stern *et al.* 2012). This need for independence is widely recognized in policy circles but is often a challenge because the parties most interested in having an intervention evaluated are often those who are most closely involved. Demonstrated independence of the evaluation process can also enhance the probability of utilization of the knowledge gained and thus boost its potential to influence policies and actions (Bamberger 2009).

Evaluation challenges

There are many challenges associated with making robust evaluations of conservation interventions, all of which are made in complex contexts (see Baylis *et al.* 2015, Ferraro 2009, Ferraro and Pattanayak 2006, Jagger *et al.* 2010, Milder *et al.* 2015, Pattanayak *et al.* 2010). It would obviously be easier if new interventions were applied as suitably designed and well-replicated randomized field trials, but that sort of experimental approach is not yet common in conservation (Jayachandran *et al.* 2016, Ferraro and Miranda 2014). For this reason, various forms of selection bias make it difficult to establish attribution or causality (*i.e.* to determine the extent to which the outcomes are due to the intervention rather than to the characteristics of units selected into the intervention). Impact evaluation is fundamentally about making this distinction by estimating the causal effect of an intervention as the difference between outcomes observed with the intervention and those that would have occurred under the ‘counterfactual’ scenario if a treated unit had not received the intervention.

Common methods to construct a counterfactual when there are concerns about selection bias are difference-in-differences (*i.e.* before-after control intervention-BACI), instrumental variables, regression discontinuity designs, matching through propensity scores (Ferraro 2009), and the most recent use of hierarchical regression models (Mitchell *et al.* 2015). Counterfactuals can also be constructed in a qualitative manner using recall or retrospective data, natural experiments, and process tracing. These advances notwithstanding, counterfactual analyses remain challenging due to their high costs in time and other resources, and they are demanding in terms of technical skills. But despite these difficulties, counterfactual thinking remains essential to reliably estimate the causal effects of conservation interventions and thereby inform policy and practical choices about whether to continue, modify, expand or drop a particular intervention (Craigie *et al.* 2015).

¹ As it is the case already for some existing initiatives, such as for example the Value and Impact Analysis (VIA) initiative of the ISEAL Alliance (<https://www.isealalliance.org/VIA>).

PREVIOUS ASSESSMENTS OF FSC IMPACTS ON NATURAL FOREST MANAGEMENT

The several available assessments of FSC's impacts are informative but most suffer from limitations, especially in their ability to attribute the observed changes to the FSC certification intervention. These studies were recently reviewed by Burivalova *et al.* (2016), an effort that will not be repeated here. What we focus on instead is the range of approaches previously used and especially on any limitations we perceive in their basis for causal inference.

Impact evaluations of FSC certification on rates of deforestation within and around FMUs based on remote sensing data are particularly numerous, perhaps because they do not require ground-based access or the cooperation of forest managers and other local stakeholders. Although remote sensing techniques could be used in many different ways in the assessment of forest management impacts (see discussion below of the studies by Lopatin *et al.* 2016, Blackman 2013), their use in the tropics to date has been restricted to deforestation assessments. While FSC certification was not explicitly designed to stop deforestation, it might nevertheless affect the probability of deforestation at the FMU level, at the level of a company that manages several FMUs, or at the jurisdictional level where the FMU is located. For remote-sensing analyses to reveal changes in forest cover in FMUs along the *certification continuum* (i.e. for firms with a range of FSC certification states), a theory-of-change or model that conveys the actions that take place with the intervention is needed to elucidate the ranges of factors associated with observed changes in forest cover. For example, lack of governmental support for enforcement of FMU rights or high costs of enforcement may constrain FMUs from stopping deforestation (Amacher *et al.* 2012, Coleman and Steed 2009, McElwee 2010). Likewise, failure of negotiations between FMU managers and local communities can create conditions that compromise the permanence of the forest (Cerutti *et al.* 2015). Conversely, some acceptable amount of deforestation by local people who reside within an FMU might result from the sorts of social welfare considerations and negotiations required for obtaining and maintaining FSC certification. In any case, deforestation outcomes clearly need to be framed within a theory-of-change that makes reference to the full set of contextual information from political economy analyses, site-specific characteristics of the FMUs, and information about the dynamics of certification (e.g. the moment when companies were certified, expect to be certified, or lost their certificate). Previous studies, which generally compared deforestation rates in certified and not certified FMUs, vary in whether they employed robust counterfactual analyses and use statistically rigorous, spatially explicit econometric approaches to evaluate whether and how FSC certification affected deforestation (e.g. Blackman 2012, Heilmayr and Lambin 2016, Miteva *et al.* 2015, Rana and Sills 2016).

One previous effort at assessing FSC certification impacts on deforestation at the country level used panel data from the Food and Agriculture Organization of the United Nations (FAO) for 1972–1994 and 2005–2010 for developed and

developing nations (Damette and Delacote 2011). Potentially explanatory variables included in the analysis were the volumes and values of harvested timber. Control variables included: national-level indicators of institutional quality (i.e. indices of the political rights and civil liberties enjoyed by local communities); a model of deforestation as a function of the country's GDP; annual GDP growth; population density in areas adjacent to the FMUs; and, the country's forest cover at the beginning of the study period. Results from this study suggest that countries in which FSC certification was prominent experienced less deforestation (Damette and Delacote 2011), but did not address the underlying causal mechanisms that resulted in this outcome.

More detailed remote-sensing analyses of deforestation were carried at country levels using spatially-explicit econometric methods in Mexico (Blackman *et al.* 2015), Indonesia (Miteva *et al.* 2015), Peru and Cameroon (Panlasigui *et al.* 2015), and Chile (Heilmayr and Lambin 2016). These analyses, which all used Hansen *et al.* (2012) global forest cover data, reported little to no effect of FSC certification on deforestation. In Indonesia, FSC certification was associated with increased “perforated” area (i.e. non-forested patches within forested ones).

In a recent study conducted in Finland, Lopatin *et al.* (2016) scrutinized the extent to which certification criteria can be assessed with remote-sensing data. They based their study on the Programme for the Endorsement of Forest Certification (PEFC, www.pefc.org), which defined 32 criteria associated with responsible forest management. The authors consolidated a list of indicators to avoid duplication and determined that remote-sensing tools are helpful at assessing outcomes related to compliance with areal limitations on harvests, respect of riparian buffer zones, and overall compliance with national regulations related to areas to be protected within FMUs (e.g. reindeer husbandry sites).

Ground-based studies on the impacts of FSC certification carried out to date vary in the extent to which they comply with robust evaluation standards. The studies in Malaysia by Imai *et al.* (2009) and in Gabon by Medjibe *et al.* (2013), for example, were each based on comparisons of a single FSC-certified concession with an adjacent non-certified concession that differed in ways that might be expected to confound interpretation of the comparisons. With a somewhat improved design, Griscom *et al.* (2014) compared logging-induced carbon emissions from three FSC-certified with six nearby non-certified concessions in Indonesia. Concessions in the two groups were similar in terms of area, forest structure, mean slope angles and elevations; logging intensity was treated as a covariate in their analyses. Although they found no discernible FSC-impact on carbon emissions, confidence in this result is reduced by the many variables unaccounted for that might have biased the results (e.g. FMU management history).

FSC-certified community forests and non-certified operations in Tanzania were compared on the basis of biodiversity as indicated by forest structure, governance, perceptions by local community members of the benefits of certification,

resource use (Kalonga *et al.* 2016), income and benefit distribution (Kalonga *et al.* 2015a), forest regeneration, and fire incidence (Kalonga *et al.* 2015b). Based on a range of mixed methods including triangulation, retrospective adaptation of a quasi-experimental approach, and a detailed value-chain analysis, the authors concluded that FSC certification was associated with enhanced conservation and development outcomes. These findings are intriguing, but causal inference is unfortunately limited in all cases because of the weak counterfactual construction. For instance, in Kalonga *et al.* (2015 a, b) the certified and non-certified operations differed in whether local people managed the forest themselves or contracted out that work; this difference renders it difficult to attribute the improved conditions to the certification intervention. In their subsequent study, Kalonga *et al.* (2016) identified forest management regime as a confounding variable, along with elevation and ease of access. They then combined the matching-paired analysis with information about perceptions of certification benefits and concluded that they could not unequivocally attribute the observed differences to the FSC certification intervention. For example, lower logging intensities in FSC areas might be due to the logging histories of each site rather than to the effect of improved management in certified operations.

In a field-based evaluation of the social impacts of FSC certification in Republic of Congo, Gabon, and Cameroon, Cerutti *et al.* (2015) compared three certified and three non-certified FMUs in each country. They reported improved social outcomes (*e.g.* improved living conditions and better benefit-sharing mechanisms) in communities adjacent to FSC-certified FMUs when compared to communities near non-certified FMUs, but the former did reportedly experience more conflicts with local people over customary rights. As noted by the authors, the observed outcomes might be due to the effect of other confounding variables, both observable and unobservable (Cerutti *et al.* 2015). FMUs were selected on the basis of several proxies that aimed to maximize similarities in traits except for certification status. These included the existence of alternative employment opportunities in the area (*i.e.* avoiding FMUs with nearby mining or agriculture activities), extent of dependence on cash crops (*i.e.* as a function of differential access to markets and thus potential causes of conflicts), ethnic characteristics that affect customary norms, administrative jurisdictions, target commercial species, and extent of set-asides. Nevertheless, important variables that define management practices were not considered (*e.g.* area and ownership of the FMU, market outlets) and control for spillover effects was difficult.

A recent field-study on the impacts of certification was carried out in Sweden (Nordén *et al.* 2016). Forest degradation was assessed with governmental forest inventory data before and several years (5–7) after logging in FSC- and PEFC-certified FMUs and in non-certified FMUs. Variables assessed correspond to management improvements expected to result from certification adoption such as extent of the environmentally important areas preserved, number of trees and high stumps remaining after felling, and compliance with set-aside requirements. Using a counterfactual comparison

through consideration of confounding variables (*e.g.* management characteristics, attributes of forest operations, and socioeconomic and biophysical information including access and location), the authors found no discernible impacts of either FSC or PEFC on the outcomes assessed, nor any differences between the certification schemes.

OUR APPROACH TO FSC CERTIFICATION AND ITS EVALUATION

How FSC certification was designed to work

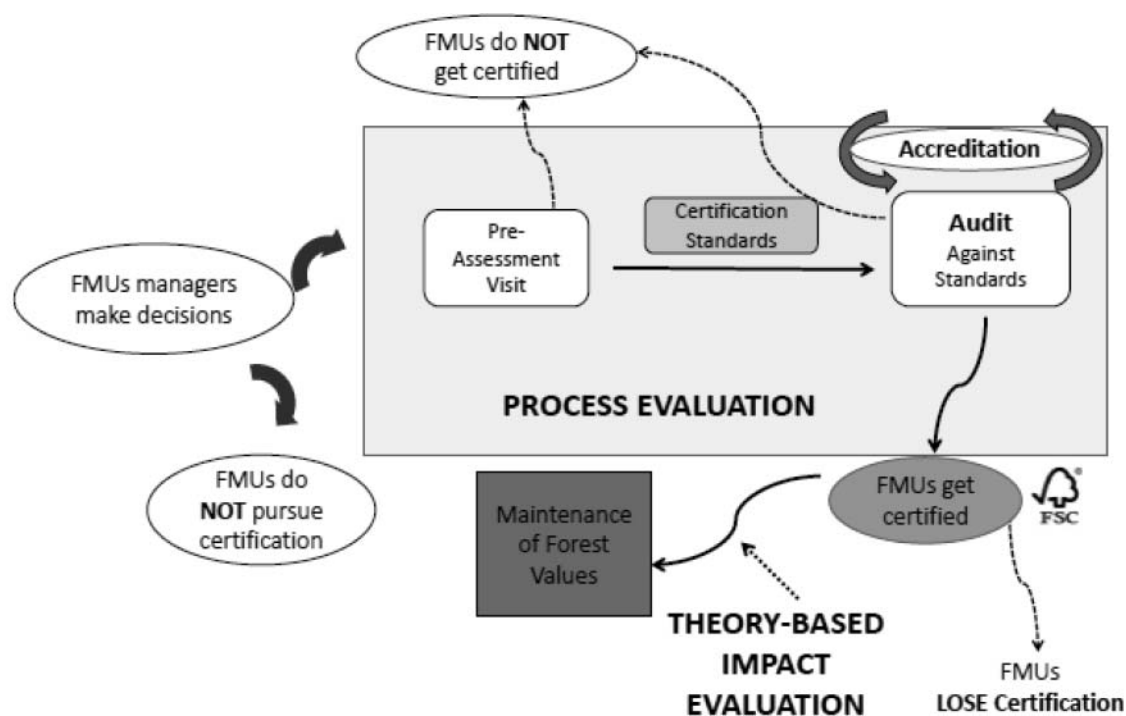
FSC certification aims to improve forest management through the adoption of better practices. FSC certification operates through a multi-stakeholder agreed-upon set of standards that are codified in ten principles along with associated criteria and indicators (www.fsc.org). Compliance with these standards is used to indicate that forest management was carried out in a responsible manner. These standards include compliance with national regulations and avoidance of illegal activities, sustained timber yields, reduced forest degradation and deforestation, maintenance of biodiversity and ecosystem services, enhanced capacities for improved management quality along the market value chain, safe and healthy working and living conditions, effective conflict avoidance and resolution mechanisms, and good relations with nearby communities (FSC 2015). In addition to the global FSC standards, some regions and countries have standards adopted to reflect the characteristics of their forests. A FSC label on a forest product, which indicates compliance with the standards, is hoped to secure market advantages such as increased market access and share, as well as price premiums.

Compliance with FSC principles is determined through the combined processes of third-party audits (*i.e.* independent verification that operations abide by FSC standards) and accreditation of the auditing process. Basically, auditors with recognized credentials and affiliation with an accredited certifying body inspect candidate FMUs to assess whether management practices meet FSC standards. Auditors inspect documents, talk with forest managers and workers, and hold open workshops to collect information about compliance with the criteria. Transparency of this process is increased by the required posting of public summaries of audit reports on FSC's website (FSC, www.fsc.org). Accreditation of certifying bodies is also implemented by a third-party organization – Accreditation Services International (ASI, www.ASI.org), which assesses the quality of the audit processes and overall performance of the FSC-approved certifying bodies.

FSC evaluation rationale and activities

One central element of an assessment FSC certification is a *process evaluation* to determine whether FSC certification was implemented according to FSC's design specifications (Figure 1). This portion of the evaluation provides insights about auditing and accreditation practices as they are implemented in the field. Process evaluation requires full

FIGURE 1 Key steps and decisions related to implementation of FSC certification. The shaded box indicates activities assessed by a **process evaluation** (e.g. auditing and accreditation). **Theory-based evaluations** assess whether, once FSC certification has been implemented as designed, the expected forest values are maintained. FMUs = forest management units, the entities that can be certified such as concessions, privately owned forests, and community forests



engagement of the parties involved in implementation of the intervention. Project partners for this type of evaluation include NGOs and other organizations locally relevant to certification including representatives of certification adoption supporting institutions (i.e. certification coaches), certifying bodies, auditors, ASI, FSC itself, and FMU managers and workers).

The second component of the evaluation is a theory-based, empirical impact evaluation, based on FSC's own theory-of-change. This component aims to assess the extent to which the FSC intervention, if implemented as designed, caused changes relative to the counterfactual condition and relative to its goals. This field-based research needs to draw on the expertise and secure the participation of a group of partners that overlaps considerably with those involved in the process evaluation and includes representatives of NGOs, FMU management and field workers, communities near the forests, local government, timber buyers, and end-product consumers.

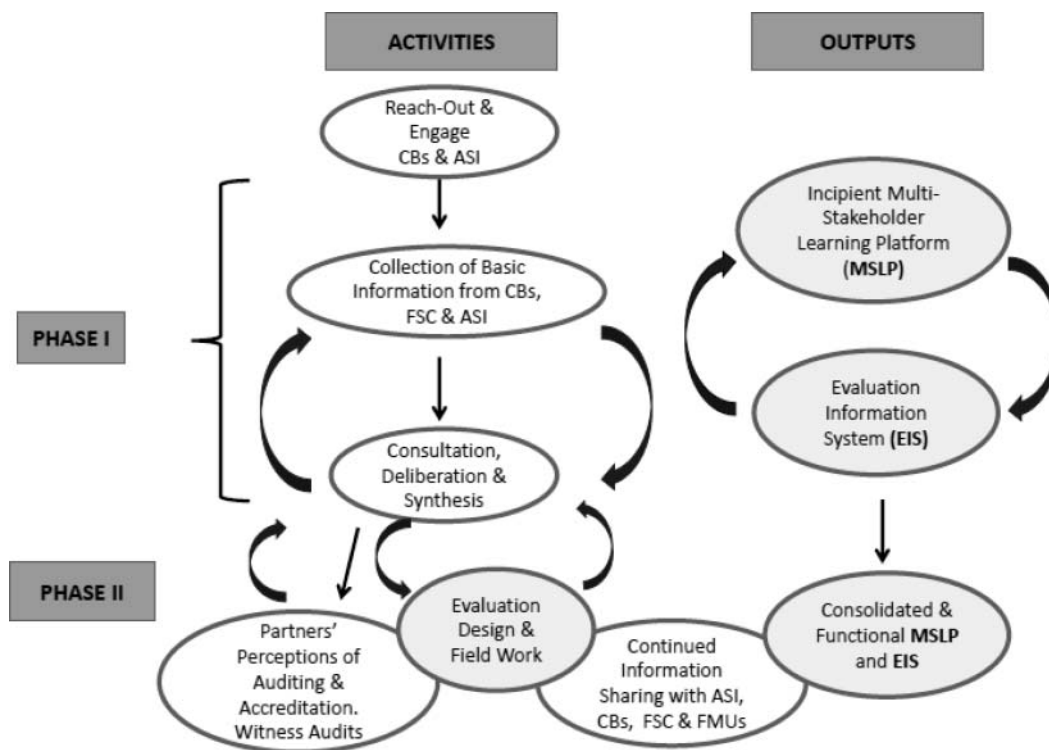
Evaluation research is facilitated by the consolidation of an active learning community that we refer to as a *Multi-Stakeholder Learning Platform* (MSLP, Figure 2). Members of this platform include representatives of organizations that helped FMUs become and remain certified, certifying bodies, auditors, FMU managers and workers, NGOs and other civil society organizations interested in forest management, consumer groups, and members of the evaluation research team. One initial outcome of this platform is an *Evaluation Information System* (EIS) that consists of the set of interested,

participating, and affected parties, arrangements (e.g. confidentiality and other non-disclosure agreements), data, and processes (e.g. workshops, field questionnaires, electronic surveys, phone interviews) through which information is collected, discussed, shared, published, and routed back into the adaptive forest management and certification decision-making processes.

A ROADMAP TOWARDS EVALUATION OF FSC CERTIFICATION

Flows of information derived from planned and on-going evaluations are needed to build trust among those involved in the evaluation process (Chatham House and UN-REDD 2011) in ways that enhance mutual social learning (Bidwell *et al.* 2013). The roadmap described here (Figure 2) is intended to inform plans for impact evaluation of certification but also to contribute to the body of knowledge about evaluation of conservation interventions. The evaluation process is based on understanding of the characteristics of the units to which the intervention was applied (i.e. FMUs), the temporal dynamics of adoption of the intervention, and the contextual factors that likely influenced the process by which particular units self-selected into and out of the intervention. Better understanding of these variables help to shape more perceptive and policy-relevant impact evaluations that, in turn, will feed back into social learning and improved decision-making.

FIGURE 2 Operational model of an evaluation of FSC certification. Deliberations and syntheses are iterated throughout the evaluation. Both the Multi-Stakeholder Learning Platform (MSLP) and the Evaluation Information System (EIS) continue to be consolidated as the evaluation progresses. New partners join the MSLP as the evaluation progresses.



Phase I of the roadmap (Figure 2) is the preparatory stage for the on-the-ground empirical evaluation (*Phase II*). The overall objective of *Phase I* is to work with a diverse group of stakeholders to gather the information required to design an impact evaluation that is credible (*i.e.* technically adequate for handling evidence), salient (*i.e.* relevant and valuable to decision-makers and other users), and legitimate (*i.e.* fair in its knowledge gathering, unbiased and respectful; Mollinga 2010, Rowe 2012). This plan should be robust enough to respond to the needs of the evaluation challenge while remaining sufficiently flexible to include the national context and its certification history and dynamics. This phase helps in the assembly of MSLP to engage others involved in evaluating certification and related interventions, and advances discussions about the roles of certification in forest conservation and its implications for local livelihoods. *Phase II* evaluation design and field-data collection are based on discussions that reveal necessary local adjustments to FSC's generic theory-of-change (FSC 2014).

FSC certification commences when a FMU voluntarily pursues that status and proceeds if audits by an accredited certifying body indicate that its management satisfies FSC standards. Both whether an FMU opts to try for FSC certification and whether it achieves that goal are both likely related to factors that also influence the outcomes of interest, which are the quality and extent of forest cover, the well-being of local populations, timber profits and their distribution, and governance issues. Thus, in order to estimate counterfactual

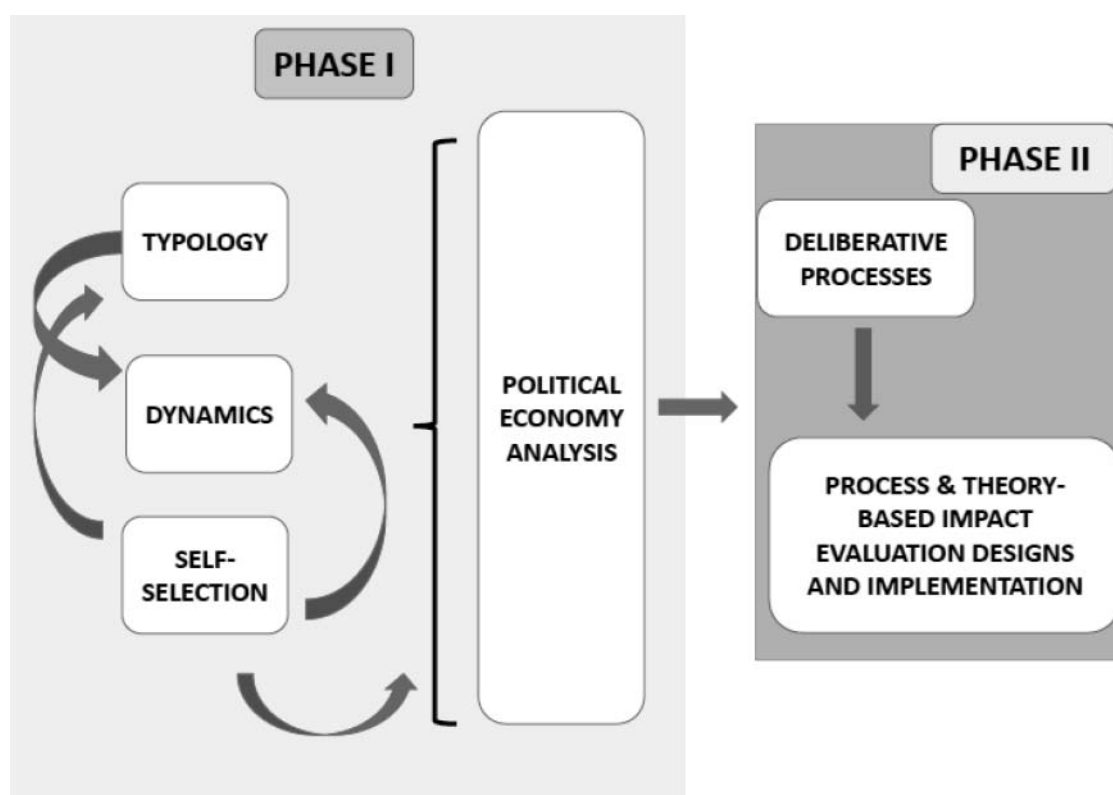
outcomes, a certified FMU cannot simply be compared with any non-certified FMU. Similarly, comparisons of conditions pre- and post-certification will not suffice as a standard for evaluation due to the impacts of political, economic, and social conditions that changed over the same period of time. In other words, both naïve comparisons (*i.e.* certified versus non-certified and pre- versus post-certification) suffer from inability to confidently attribute to the certification intervention the differences or changes observed. Identification of non-certified FMUs that can be employed to reveal the counterfactual outcomes for certified FMUs requires detailed understanding of the potential confounding factors that influence both certification and the outcomes of interest.

REQUISITE BACKGROUND STUDIES

For each country where the impacts of FSC are to be evaluated, background studies should include a political economy analysis of the timber and related sectors where the intervention is to be evaluated, FMU typologies, temporal dynamics of FSC certification, and an analysis of factors that determine FMU selection into certification (Figure 3). This knowledge needs to be updated regularly given the ever-present dynamism.

The political economy analyses that undergird evaluations of FSC certification characterize the contextual factors that determine the fates of FMUs such as actions that pertain to land cover change and the concept of territory, its zoning, and

FIGURE 3 Roadmap for evaluation of the biophysical, social, economic, and policy impacts of FSC certification. **Phase I studies** use insights from a variety of stakeholders to identify key elements of process-focused and theory-based evaluations of the impacts of FSC certification. **Phase II** is based on FSC's theory-of-change, and involves formulation of research hypotheses, sampling design, field data collection and analyses, and consolidation and sharing of knowledge gained.



uses. Because natural forests and their management are framed within the frequently changing and all-encompassing political and socio-economic settings in often-contested territories, the forests are used for different purposes and, as such, suffer a range of fates. Precursors for the proposed background studies on political economy as related to certification include work by Ruiz-Pérez and colleagues (2005) in the Congo Basin and Salazar and Gretzinger (2005) in Central America.

Political economy analyses should capture an understanding of the characteristics of the forest sector that are relevant to FSC certification. They describe the historical, political, social, and economic contexts in which certain decisions determined the fates of managed natural forests. They also provide timelines of key events and processes expected to have influenced choices made vis-à-vis certification (e.g. the fall of President Soeharto in Indonesia and decentralization in both Indonesia and Brazil) including policies and legal frameworks, their evolution through time, and how other social and economic processes influenced land-use planning and land-cover change. For example, whether an FMU is located on an active forest frontier or in a remote region will likely influence both the outcomes of certification in particular and forest management in general. More generally, decisions relevant to forest management are made at a range of political levels that vary with the mode of governance and institutional regimes as well as in response to policy shifts

(Coleman 2009) and individual preferences. Political economy background studies for certification impact research also need to account for the emergence of FSC in each country and how the FSC influenced the evolution of national and even broader regulatory frameworks. The most direct of these influences were on national certification schemes (e.g. LEI in Indonesia, CERFLOR in Brazil), legality certification (e.g., SVLK in Indonesia) and voluntary partner agreements through the Forest Law Enforcement, Governance & Trade (FLEGT) in a range of tropical countries. The existence of these influences is recognized (Cashore and Stone 2012, Overdevest 2010) and their on-the-ground relevance awaits empirical demonstration (e.g., Robalino *et al.* 2015 about synergies of PES with other interventions).

A deeper understanding of FMUs is provided by typological analyses based on their characteristics. This analysis provides a systematic way to appreciate similarities and differences among FMUs based on attributes that likely influenced their probability of becoming certified as well as the expected outcomes of certification. That is, typologies help to distill the relationships among groups of variables and overall differences among FMUs. Specifically, a typology describes groups of FMUs that are similar in operational and institutional characteristics as well as in their geographies and stage along the *certification continuum* (Romero *et al.* 2013). Typologies can help guide the selection of counterfactuals.

A third important background study recognizes that FMUs are located along a *certification continuum* of forest management practices that represent the stages in the certification process (e.g. from never having engaged in FSC certification to having remained certified for several years or having lost certification; see Romero *et al.* 2013). Classification of FMUs into the simple categories of “certified” and “non-certified” fails to capture this complexity. And because involvement in certification is an on-going process and not a one-time choice, decisions made over time by FMUs along the *certification continuum* need to be documented and related to the nature of FSC implementation processes in a given country. In doing so, particular windows of time when contextual factors likely either facilitated or obstructed engagement in certification can be identified, and more generally, patterns of decisions made vis-à-vis certification by FMU managers at particular times can be identified. These decisions are subject to change in response to a variety of political, economic, and other factors related to the timber and associated sectors (e.g. investments, competing opportunities, market realities, changing legal frameworks), all collected in the political economy analysis. In particular, contextual factors that operate at local, national, and international levels can influence FMU decisions about whether to opt for certification and, once certified, whether to remain certified. At the same time, market dynamics (e.g. consumer preferences and acquisition power) change and influence suppliers’ decisions vis-à-vis certification. Shifting legal frameworks and their enforcement, changes in certification standards, novel technical capacities, technological innovations, global/regional/national economic conditions, availability of external support, and perceived and realized cost-benefit ratios are among the factors that can affect FMU decisions about certification (Chen *et al.* 2011, Crow and Danks 2010, Kollert and Lagan 2007, Nebel *et al.* 2005, Vincent 2016).

The fourth type of study needed to reveal FSC impacts deals with the motivation behind choices made by FMUs. Although certification was initially conceived of as a market mechanism that would provide “green” premiums to firms that satisfy auditors that they use specified management practices, firms may choose certification for a wide variety of non-fiscal reasons. In any case, the private benefits of forest certification actually realized by firms are not consistent across time scales, products, regions, countries, and type of firms (Blackman and Guerrero 2012). In addition to expected improvements in management practices, certified firms reportedly have enhanced learning and transparency, increased public confidence and social acceptance, social welfare improvement, and greater environmental responsibility (Araujo *et al.* 2009, Cabbage *et al.* 2010, Vidal and Kozak 2008). A review of available literature on certification decisions along with analyses of trends of the certification dynamics studies discussed previously helped to structure the methodology used for the proposed study of self-selection into certification. Self-selection can be explored with semi-structured interviews and informal conversations with a range of social actors with stakes in FSC certification.

IMPEDIMENTS TO ROBUST EVALUATIONS OF THE IMPACTS OF FSC CERTIFICATION

There are various reasons why FSC remains to be evaluated rigorously with independently collected field data. First of all, rigorous evaluation was not planned for when the intervention was designed, which means that baseline data were not collected before its implementation and specification of the mechanisms of change and factors that might affect self-selection were not considered (Craigie *et al.* 2015). Even with perfect foresight, it would remain difficult to disentangle the effects of FSC certification from many potential confounding factors. The counterfactual method approach to impact evaluation, which is relatively new in conservation, can help elucidate impacts, but it is challenging to implement well (Andam *et al.* 2008, Ferraro *et al.* 2011, Miteva *et al.* 2012, Pfaff and Robalino 2012, Pfaff *et al.* 2013). Second, lack of on-the-ground impact evaluation of FSC might stem from an understandable aversion to possible negative outcomes (Bamberger and Kirk 2009, Meek *et al.* 2015). For donors already weary of investing in the tropical forestry sector, paying the substantial cost of an evaluation of certification may not be very attractive (Craigie *et al.* 2015). Given the pervasiveness of bad management in tropical forestry, including illegal harvests and lack of incentives to adopt the practices needed to sustain timber yields, donor fatigue is understandable. There might also be the feeling that it is less important to fine-tune a complex and already functioning intervention than it is to address a seemingly simpler problem such as legality (e.g. Voluntary Partnership Agreements within the EU’s FLEGT Action Plan), or other interventions with more simple outcomes to be measured (e.g. avoided deforestation in protected areas). Advocates for certification within donor agencies may also be reluctant to expose the intervention that they have championed to the rigors of independent evaluation.

Potential financial backers for field-based evaluations of certification might also wonder why they should fund studies on an intervention that is obviously struggling for market share, when “hot” new approaches seem almost certain to deliver some of the promised benefits (e.g. zero deforestation pledges). While proponents of forest certification in general and the FSC in particular actively seek to expand and improve the impacts of the intervention they endorse, over the decades since certification started, new approaches to conservation and development have been introduced (e.g. PES, REDD+). Each in the series of tropical forest conservation strategies was initially promoted as a “silver bullet,” by which is meant that they were each marketed as “perfect” or “win-win” ways to attack the problem. These interventions, including forest certification, followed a familiar trajectory after first being introduced with fanfare by designers, embraced enthusiastically by some influential scholars, supported by donors, adopted by fund-seeking conservation practitioners, investigated by researchers, and finally pronounced *passé* by pundits (Lund *et al.* 2016).

Hope for silver bullets does not seem to diminish despite the fact that conservation problems are complex and prone to morph over time and vary over space, which renders universal

and simple solutions patently unlikely. Market-based interventions, like certification, seem particularly prone to rejection by some critics even without the insights derived from robust impact evaluations (*e.g.* Fletcher *et al.* 2016) and even while the interventions proclaimed *passé* continue to attract new proponents and expand in area covered. Although earlier approaches might be trivialized as simply passing fads (Redford *et al.* 2013), they are not even passed. Ecotourism did not save the tropics and most non-timber forest product harvests are not as lucrative as once portrayed, but both continue to contribute substantially to many local conservation and development initiatives. Another example is the much-criticized integrated conservation and development project (ICDP) approach (*e.g.* Wells and Brandon 1997); it may currently be hard to find funding for an ICDP that self-identifies as such, but many ICDP components are evident in REDD+ projects (Bauch *et al.* 2014). The list goes on, but the relevance of this discussion here is that certification, as a market-based instrument to promote conservation and development, has also been relegated to the trash heap of failed initiatives by some critics. To the extent that donors, research agencies, and trendy researchers hear and accept this proclamation, perhaps it is no surprise that investments in the evaluation of certification have so far not materialized. Instead, FSC certification, which is an established component of a diverse conservation portfolio, is disregarded in favor of a focus on new initiatives such as investor-driven approaches and corporate pledges of zero deforestation.

Rigorous, independent, empirical, field-based evaluations of the impacts of FSC certification will require substantial input from FSC, certifiers, and those NGOs that have strongly supported FSC, all of whom have vested interests in the intervention and justifiable concerns about the outcomes of the evaluations. The people engaged in certification are also busy and they, and the organizations for which they work, recognize that they are unlikely to be financially compensated for the time they invest in evaluation of an intervention that they feel certain is working (*i.e.* lack of incentives for evaluation, Craigie *et al.* 2015). On the other hand, they recognize that poorly designed or otherwise weak evaluations are unlikely to contribute much to improvement of the intervention, but can be damaging if they come up with the “wrong” answer.

Until a major supporter for the adoption of FSC makes future funding contingent upon its evaluation and makes funding available for that purpose, the many insights that would emerge from such an endeavor will remain out of reach.

CONCLUSIONS

The collection and synthesis of information needed to construct a credible field-based evaluation of FSC’s impacts on natural forest management in the tropics is an interdisciplinary endeavor that necessarily involves stakeholders who have not previously worked together (*e.g.* Palmer *et al.* 2016). Evaluations of complex, large scale, and long-term conservation interventions like FSC certification will also be expensive because, if well-done, they will require participation of

many stakeholders, extended time in the field by well-informed observers, and substantial buy-in by governments. In the case of forest certification, for example, it may be relatively easy for researchers to gain access to FSC-certified FMUs and those that seek to become so, but substantial leverage may be needed to gain equal access to never certified FMUs and those that have lost certification. Despite these and other challenges, approaches like the one we describe may provide insights about certification’s added values, suggest ways to improve the intervention, and reveal synergies with other interventions (Robalino *et al.* 2015, Heilmayr and Lambin 2016).

A thorough evaluation of the complex intervention of tropical forest management certification by the FSC can be expected to yield complicated and heavily nuanced answers. First of all, comparisons of FSC-certified FMUs with carefully selected (or constructed) counterfactuals as well as against FSC’s stated goals (*e.g.* sustained timber yields, safe working conditions) need to be outcome-by-outcome. It might be tempting to calculate a composite score, but it is unclear how the different effects should be weighted (and an unweighted approach simply assumes equal weights). Combining metrics for worker safety with biodiversity retention and stream crossings seems like an exercise in obfuscation. Even with the thorough background research we describe, it will be challenging to differentiate between FSC’s direct and indirect impacts. For example, the three-pillared structure of the FSC, with its social, business, and environmental chambers, as well as its equal representation of the global “north” and “south,” are much-mimicked. Similarly, although the FSC was certainly not the first to initiate third-party audits, by doing so it influenced the design of other conservation interventions. Even more broadly, the FSC helped gain credence for conservation based on sustainable forest management, which will be a challenging impact to measure.

Finally, if it is ultimately not possible to separate the effects of FSC certification from those of other contributing factors, then claims about FSC’s successes should be considered with caution. Ignoring this limitation would be a disservice to the conservation community and may undermine the use of the certification tool. Awareness of the difficulties of learning from FSC certification implementation through impact evaluation would also encourage design of interventions that avoid some of the attribution analysis bottlenecks while promoting adaptive management and sustainability.

ACKNOWLEDGMENTS

This paper is based on a series of baseline studies prepared for a CIFOR project entitled: ‘*Evaluation of the Impacts of FSC Certification of Natural Forest Management in Tropical Forests.*’ We thank all the participants in this project and particularly the representatives of certification coach organizations, training institutions, certification bodies, and forest management units in Indonesia, Peru, Brazil, and Gabon, who shared with us their reflections on forest certification. We also

appreciate the comments of four anonymous reviewers. This study was made possible by funding from the United States Agency for International Development (USAID) and from the CGIAR Program on Forests, Trees and Agroforestry. We also thank PROFOR-World Bank for the initial support without which many of the ideas in this paper would not have emerged. A presentation of this study at the Central African Forests and Institutions Workshop on Central African Forests in Paris in September 2013 evoked many useful comments.

REFERENCES

- AMACHER, G.S., OLLIKAINEN, M., and KOSKELA, E. 2012. Corruption and forest concessions. *Journal of Environmental Economics and Management* **63**: 92–104.
- ANDAM, K.S., FERRARO, P.J., PFAFF, A., SANCHEZ-AZOFEIFA, G.A., and ROBALINO, J.A. 2008. Measuring the effectiveness of protected area networks in reducing deforestation. *Proceedings of the National Academy of Sciences of the United States of America* **105**: 16089–16094.
- ARAUJO, M., KANT, S., and COUTO, L. 2009. Why Brazilian companies are certifying their forests? *Forest Policy and Economics* **11**: 579–585.
- BAMBERGER, M., RUGH, J., and MABRY, L. 2012. *Real World Evaluation: working under budget, time, data, and political constraints*. Sage Publications. 666 pp.
- BAMBERGER, M., and KIRK. 2009. *Making Smart Policy: Using Impact Evaluation for Policy Making Case Studies on Evaluations that Influenced Policy*. Doing Impact Evaluation Series #14. World Bank. Washington, DC. 75 pp.
- BAMBERGER, M., CARDEN, F., and RUGH, J. 2009. Alternatives to the conventional counterfactual. *Real World Evaluation: Summary of Session 713 Think Tank*. American Evaluation Association, Orlando. [Available at: <http://www.realworldevaluation.org>; accessed June 2016].
- BAUCH, S.C., SILLS, E.O., and PATTANAYAK, S.L. 2014. Have we managed to integrate conservation and development? ICDP impacts in the Brazilian Amazon. *World Development* **64**: S135–S148.
- BAYLIS, K., HONEY-ROSES, J., BORNER, J., CORBERA, E., EZZINE-DE BLAS, D., FERRARO, P.J., LAPEYRE, R., PERSSON, U.R., PFAFF, A., and WUNDER, S. 2015. Mainstreaming impact evaluation in nature conservation. *Conservation Letters*. <http://dx.doi.org/10.1111/conl.12180>.
- BELL, S., SHAW, B., and BOAZ, A. 2011. Real-world approaches to assessing the impact of environmental research on policy. *Research Evaluation* **20**: 227–237. DOI: 10.3152/095820211X13118583635792
- BIDWELL, D., DIETZ, T., and SCAVIA, D. 2013. Fostering knowledge networks for climate adaptation. *Nature Climate Change* **3**: 610–611.
- BLACKMAN, A., GOFF, L., and RIVERA PLANTER, M. 2015. Does eco-certification stem tropical deforestation? Forest Stewardship Council certification in Mexico. *Resources for the Future Discussion Paper* 15–36. Washington, DC.
- BLACKMAN, A. 2013. Evaluating forest conservation policies in developing countries using remote sensing data: an introduction and practical guide. *For. Policy Econ.* **34**: 1–16.
- BLACKMAN, A. 2012. *Ex post* evaluation of forest conservation policies using remote sensing data. An introduction and practical guide. Resources for the Future, Discussion Paper 12-05.
- BLACKMAN, A., and GUERRERO, S. 2012. What drives voluntary eco-certification in Mexico? *Journal of Comparative Economics* **40**: 256–268.
- BURIVALOVA, Z., F. HUA1, L.P. KOH, C. GARCIA, and F.E. PUTZ. 2016. A critical comparison of conventional, certified, and community management of tropical forests in terms of environmental, economic, and social variables. *Conservation Letters* doi: 10.1111/conl.12244.
- CASHORE, B., and STONE, M.W. 2012. Can legality verification rescue global forest governance? Analyzing the potential of public and private policy intersection to ameliorate forest challenges in Southeast Asia. *Forest Policy and Economics* **18**: 13–22.
- CERUTTI, P.O., LESCUYER, G., TSANGA, R., KASSA, S.N., MAPANGOU, P.R., MENDOULA, E.E., MISSAMBALOLA, A.P., NASI, R., ECKEBIL, P.P.T., and YEMBE, R.Y. 2014. Social impacts of the Forest Stewardship Council certification: An assessment in the Congo basin. Occasional Paper 103. CIFOR, Bogor, Indonesia. 74 pp.
- CHAPLIN-KRAMER, R., JONELL, M., GUERRY, A., LAMBIN, E.F., MORGAN, A.J., PENNINGTON, D., SMITH, N., ATKINS FRANCH, J., and POLASKY, S. 2015. Ecosystem service information to benefit sustainability standards for commodity supply chains. *Annals of the New York Academy of Sciences* **1355**: 77–97.
- CHATHAM HOUSE AND UN-REDD. 2011. *Guidance for the Provision of Information on REDD+ Governance (Draft)*. London and Geneva. 42 pp (available at <http://theredddesk.org/resources/draft-guidance-provisioninformation-redd-governance>; accessed June 2016).
- CHEN, J., TIKINA, A., KOZAK, R., INNES, J., DUINKER, P., and LARSON, B. 2011. The efficacy of forest certification: perceptions of Canadian forest products retailers. *The Forestry Chronicle* **87**: 636–643.
- COLCHESTER, C., and LOHMAN, L. 1990. *The Tropical Forestry Action Plan: What Progress?* The World Rainforest Movement. Third World. 104 pp.
- COLEMAN, E.A. 2009. Institutional factors affecting ecological outcomes in forest management. *Journal of Policy Analysis and Management* **28**: 122–146.
- COLEMAN, E.A., and STEED, B.C. 2009. Monitoring and sanctioning in the commons: an application to forestry. *Ecological Economics* **68**: 2106–2113.
- CRAIGIE I.D., BARNES, M.D., GELDMANN, J., and WOODLEY, S. 2015. International funding agencies: potential leaders of impact evaluation in protected areas? *Philosophical Transactions of the Royal Society B*. **370**: 20140283.
- CROW, S., and DANKS, C. 2010. Why certify? Motivations, outcomes and the importance of facilitating organizations

- in certification of community-based forestry initiatives. *Small Scale Forestry* **9**: 195–211.
- CUBBAGE, F., DIAZ, D., and YAPURA, P. 2010. Impacts of forest management certification in Argentina and Chile. *Forest Policy and Economics* **12**: 497–504.
- DAMETTE, O., and DELACOTE, P. 2011. Unsustainable timber harvesting, deforestation and the role of certification. *Ecological Economics* **70**: 1211–1219.
- EKLIN, K., EVENSMO, I.F., GEORGESCU, I., HUBERT, V., LE, J., MALIK, T., TREYER, S., and BRUN, M. 2014. *The Committee on World Food Security Reform: Impacts on Global Governance of Food Security*. Agriculture Policy Brief 3. Paris: Institute for Sustainable Development and International Relations.
- EZZINE DE BLAS, D., WUNDER, S., RUIZ-PÉREZ, M., and MORENO-SÁNCHEZ, R. DEL P. 2016. Global patterns in the implementation of payments for environmental services. *PLoS ONE* **11**: e0149847. doi:10.1371/journal.pone.0149847
- FARLEY, K., LUCAS, S., MOLYNEAUX, J., and PENN, K. 2012. *Impact Evaluations of Agriculture Projects. Millennium Challenge Corporation. Principles into Practice*. Washington, D.C. 33 pp.
- FERRARO, P.J., and PRESSEY, R.L. 2015. Measuring the difference made by conservation initiatives: protected areas and their environmental and social impacts. *Phil. Trans. R. Soc. B* **370**: 20140270.
- FERRARO, P.J., and MIRANDA, J.J. 2014. The performance of non-experimental designs in the evaluation of environmental programs: a design-replication study using a large-scale randomized experiment as a benchmark. *Journal of Economic Behavior Organizations* **107**: 344–365.
- FERRARO, P.J., HANAUER, M.M., and SIMS, K.R.E. 2011. Conditions associated with protected area success in conservation and poverty reduction. *Proceedings of the National Academy of Sciences* **108**: 13913–13918.
- FERRARO, P. 2009. Counterfactual thinking and impact evaluation in environmental policy. *New Directions for Evaluation* **122**: 75–84.
- FERRARO, P.J., and PATTANAYAK, S. 2006. Money for nothing? A call for empirical evaluation of biodiversity conservation investments. *PLoS Biology* **4**: 482–488.
- FLETCHER, R., DRESSLER, W., BÜCHER, W., and ANDERSON, C.A. 2016. Questioning REDD+ and the future of market-based conservation. *Conservation Biology* **30**: 673–675.
- FSC–FOREST STEWARDSHIP COUNCIL. 2014. FSC’S “Theory Of Change”, Intended Impacts And Related Indicators. [Available at: ToC <https://ic.fsc.org/en/our-impact/program-areas/monitoring-and-evaluation/fsc-theory-of-change>; accessed June 2016].
- GERTLER, P.J., MARTINEZ, S., PREMAND, P., RAWLINGS, L.B., and VERMEERSCH, C.M.J. 2011. *Impact evaluation in practice*. World Bank, Washington, DC. http://siteresources.worldbank.org/EXTTHDOFFICE/Resources/5485726-1295455628620/Impact_Evaluation_in_Practice.pdf. 266 pp.
- GOVERNMENT ACCOUNTABILITY OFFICE (GAO). 2009. *Program Evaluation: a Variety of Rigorous Methods can Help Identify Effective Interventions in United States*. Government Accountability Office, Washington, DC. 49 pp.
- GRISCOM, B., ELLIS, P., and PUTZ, F.E. 2014. Carbon emissions performance of commercial logging in East Kalimantan, Indonesia. *Global Change Biology*. doi:10.1111/gcb.12386
- HANSEN, M.C., POTAPOV, P.V., MOORE, R., HANCHER TURUBANOVA, S.A., TYUKAVINA, A., THAU, D., STEHMAN, S.V., GOETZ, S.J., LOVELAND, T.R., KOMMAREDDY, A., EGOROV, A., CHINI, JUSTICE, C.O., and TOWNSHEND, J.R.G. 2013. High-Resolution Global Maps of 21st-Century Forest Cover Change. *Science* **342**: 850–853.
- HEILMAYR, and LAMBIN, E.F. 2016. Impacts of non-state, market-driven governance on Chilean forests. *Proceedings of the National Academy of Sciences* **113**: 2910–2915.
- IMAI, N., SAMEJIMA, H., LANGNER, A., ONG, R.C., KITA, S., TITIN, J., CHUNG, A.Y.C., LAGAN, P., LEE, Y.F., and KITAYAMA, K. 2009. Co-benefits of sustainable forest management in biodiversity conservation and carbon sequestration. *PLoS ONE* **4**(12): e8267. doi:10.1371/journal.pone.0008267
- JAGGER, P., SILLS, E.O., LAWLOR, K., and SUNDERLIN, W.D. 2010 *A Guide To Learning About Livelihoods Impacts Of Redd+ Projects*. CIFOR Occasional Paper 56. CIFOR, Bogor, Indonesia. 110 pp.
- KALONGA, S.K., MIDTGAARD, F., and KLANDERUD, K. 2016. Forest certification as a policy option in conserving biodiversity: An empirical study of forest management in Tanzania. *Forest Ecology and Management* **361**: 1–12.
- KALONGA, S., KULINDWA, K., and MSHALE, B. 2015a. Equity in distribution of proceeds from forest products from certified community-based forest management in Kilwa District, Tanzania. *Small-Scale Forestry* **14**: 73–89.
- KALONGA, S.K., MIDTGAARD, F.M., and EID, T.H. 2015b. Does forest certification enhance forest structure? Empirical evidence from certified community-based forest management in Kilwa District, Tanzania. *International Forestry Review* **17**: 182–194.
- KOLLERT, W., and LAGAN, P. 2007. Do certified tropical logs fetch a market premium? A comparative price analysis from Sabah, Malaysia. *Forest Policy and Economics* **9**: 862–868.
- LAMBIN, E.F., MEYFROIDT, O., RUEDA, X., BLACKMAN, A., BÖRNER, J., CERUTTI, P.O., DIETSCH, T., JUNGSMANN, L., LAMARQUE, P., LISTER, J., WALKER, N.F., and WUNDER, S. 2014. Effectiveness and synergies of policy instruments for land use governance in tropical regions. *Global Environmental Change* **28**: 129–140.
- LE VELLY, G., and DUTILLY, C. 2016. Evaluating payments for environmental services: Methodological challenges. *PLoS ONE* **11**: e0149374. doi:10.1371/journal.pone.0149374
- LOPATIN, E., THRISHKIN, M., and GAVRILOVA, O. 2016. Assessment of Compliance with PEFC Forest Certification Indicators with Remote Sensing. *Forests* **7**: 85. doi:10.3390/f7040085

- LUND, J.F., SUNGUSIA, E., MABELE, M.B., and SCHEBA, A., 2016. Promising change, delivering continuity: REDD+ as conservation fad. *World Dev.* <http://dx.doi.org/10.1016/j.worlddev.2016.08.005>.
- MASCIA, M.B., PAILLER, S., THIEME, M.L., ROWE, A., BOTTRILL, M.C., DANIELSEN, F., GELDMANN, J., NAIDOO, R., PULLIN, A.S., and BURGESS, N.D. 2014. Commonalities and complementarities among approaches to conservation monitoring and evaluation. *Biological Conservation* **169**: 258–267.
- McELWEE, P.D. 2010. Resource use among rural agricultural households near protected areas in Vietnam: the social costs of conservation and implications for enforcement. *Environmental Management* **45**: 113–131.
- McKINNON, M.C., MASCIA, M.B., YANG, W., TURNER, W.R., and BONHAM, C.(A). 2015. Impact evaluation to communicate and improve conservation non-governmental organization performance: the case of Conservation International. *Phil. Trans. R. Soc. B* **370**: 20140282.
- McKINNON, M.C., COLCHESTER, G.S.H., GARSIDE, R., MASUDA, Y.J., and MILLER, D.C. (b). 2015. Map the evidence. *Nature* **528**: 185–187.
- MEDJIBE, V.P., PUTZ, F.E., and ROMERO, C. 2013. Certified and uncertified logging concessions compared in Gabon: changes in stand structure, tree species, and biomass. *Environmental Management* **51**: 524–540.
- MEEK, M.H. 2015. Fear of failure in conservation: the problem and potential solutions to aid conservation of extremely small populations. *Biological Conservation* **184**: 209–217.
- MILDER, J.C., ARBUTHNOT, M., BLACKMAN, A., BROOKS, S.E., GIOVANNUCCI, D., GROSS, L., KENNEDY, E.T., KOMIVES, K., LAMBIN, E.F., LEE, A., MEYER, D., NEWTON, P., PHALAN, B., SCHROTH, G., SEMROC, B., VAN RIKXOORT, H., and ZRUST, M. 2014. An agenda for assessing and improving conservation impacts of sustainability standards in tropical agriculture. *Conservation Biology* **29**: 309–320.
- MITCHELL, S., GELMAN, A., ROSS, R., HUYNH, U., MCCLELLAN, L., HARRIS, M., BARI, S., CHEN, J., OHMENG-DAPAAH, S., NAMAKULA, P., EHRlich SACHS, S., PALM, C., and SACHS, J.D. 2015. The Millennium Villages Project: A protocol for the final evaluation [Available at http://millenniumvillages.org/wp-content/uploads/2015/07/MVP_paper_Lancet_FINAL.pdf; downloaded on October 22 2016].
- MITEVA, D.A., PATTANAYAK, S.K., and FERRARO, P.J. 2012. Evaluation of biodiversity policy instruments: what works and what doesn't? *Oxford Review of Economic Policy* **28**: 69–92.
- MITEVA, D.S., LOUCKS, C., and PATTANAYAK, S. 2015. Social and environmental impacts of forest management certification in Indonesia. *PLoS One* **10**: e0129675. doi:10.1371/journal.pone.0129675.
- MOLLINGA, P.P. 2010. Boundary work and the complexity of natural resources management. *Crop Science* **50**: S1–S9.
- NEBEL, G., QUEVEDO, L., JACOBSEN, J.B., and HELLES, F. 2005. Developing and economic significance of forest certification: the case of FSC in Bolivia. *Forest Policy and Economics* **7**: 175–186.
- NASON, E., KLAUTZER, L., RUBIN, J., HANNEY, S., WOODING, S., and GRAN, J. 2007. *Policy and Practice Impacts of Research Funded by the Economic and Social Research Council: A Case Study of the Future of Work Programme, Supporting Data*. Technical Report RAND Europe. Pp. 1–159.
- NELSON, A., and CHOMITZ, K.M. 2011. Effectiveness of strict vs. multiple use protected areas in reducing tropical forest fires: a global analysis using matching methods. *PLoS One* **6**: e22722.
- ORGANISATION FOR ECONOMIC CO-OPERATION AND DEVELOPMENT (OECD)/DEVELOPMENT ASSISTANCE COMMITTEE (DAC). 2002. *Glossary of key terms in evaluation and results based management*. The DAC Working Party on Aid Evaluation, OECD, Paris. www.oecd.org/dataoecd/29/21/2754804.pdf
- OVERDEVEST, C. 2010. Comparing forest certification schemes: the case of ratcheting standards in the forest sector. *Socio-Economic Review* **8**: 47–76.
- PALMER, M.A., KRAMER, J.G., BOYD, J., and HAWTHORNE, D. 2016. Practices for facilitating interdisciplinary synthetic research: the National Socio-Environmental Synthesis Center (SESYNC). *Current Opinion in Environmental Sustainability* **19**: 111–122.
- PANLASIGUI, S., RICO-STAFFRON, J., SWENSON, J., LOUCKS, C.J., and PFAFF, A. 2015. Early Days in the Certification of Logging Concessions: estimating FSC's deforestation impact in Peru & Cameroon (Nicholas Institute, Duke University, NC (Working Paper Draft).
- PATTANAYAK, S.K., WUNDER, S., and FERRARO, P.J. 2010. Show me the money: Do payments supply environmental services in developing countries? *Review of Environmental Economics and Policy* **4**: 254–274.
- PERRIN, N. 2012. *Linking Monitoring and Evaluation to Impact Evaluation*. Impact Evaluation Notes #2. InterAction: A United Voice for Global People and Rockefeller Foundation. 22 pp.
- PFAFF, A., AMACHER, G., S., and SILLS, E.O. 2013. Realistic REDD: Improving the forest impacts of domestic policies in different settings. *Review of Environmental Economics and Policy* **7**: 114–135. doi:10.1093/reep/res023
- PFAFF, A., and ROBALINO, J. 2012. Protecting forests, biodiversity, and the climate: predicting policy impact to improve policy choice. *Oxford Review of Economic Policy* **28**: 164–179.
- PROFOR and FAO. 2011. *Framework for Assessing and Monitoring Forest Governance*. Program on Forests – PROFOR and Food and Agriculture Organization-FAO. Washington and Rome. 36 pp.
- PUTZ, F.E., DYKSTRA, D.P., and HEINRICH, R. 2000. Why poor logging practices persist in the tropics. *Conservation Biology* **14**: 951–956.
- RANA, P., and SILLS, E.O. (In preparation). FSC certification of natural forest management in tropical countries: case studies in Brazil, Indonesia, and Gabon using the Synthetic Control Method. CIFOR Occasional Paper.

- REDFORD, K., PADOCH, C., and SUNDERLAND, T. 2013. Fads, funding, and forgetting in three decades of conservation. *Conservation Biology* **27**: 437.
- RESNICK, D., BABU, S., HAGGBLADE, S., HENDRIKS, S., and MATHER, D. 2015. Conceptualizing Drivers of Policy Change in Agriculture, Nutrition, and Food Security. IFPRI Discussions Paper 01414. Washington, D.C. 56 pp.
- ROBALINO, J., SANDOVAL, C., BARTON, D.N., CHACON, A., and PFAFF, A. 2015. Evaluating interactions of forest conservation policies on avoided deforestation. *PLoS ONE* **10**: e0124910. [doi:10.1371/journal.pone.0124910](https://doi.org/10.1371/journal.pone.0124910).
- ROGERS, P. 2012. *Introduction to Impact Evaluation. Impact Evaluation*. Notes #1. InterAction: A United Voice for Global People and Rockefeller Foundation, New York. 21 pp.
- ROMERO, C., and CASTREN, T. 2013. *Approaches to Measuring The Conservation Impact of Forest Management Certification*. Working Paper. Program on Forests (PROFOR), Washington, DC. 42 pp.
- ROMERO, C., PUTZ, F.E., GUARIGUATA, M.R., SILLS, E.O., and CERUTTI, P.O. 2013. *A review of current knowledge about the impacts of forest management certification and a proposed framework for its formal evaluation*. Occasional Paper 91. CIFOR, Bogor, Indonesia. 36 pp.
- ROWE, A.E. 2012. Evaluation of natural resource interventions. *American Journal of Evaluation* **33**: 384–394.
- RUIZ PEREZ, M., EZZINE DE BLAS, D., NASI, R., SAYER, J.A., SASSEN, M., ANGOUE, C., GAMI, N., NDOYE, O., NGONO, G., NGUINGUIRI, J.-C., NZALA, D., TOIRAMBE, B., and YALIBANDA, Y. 2005. Logging in the Congo Basin: a multi-country characterization of timber companies. *Forest Ecology and Management* **214**: 221–236.
- SALAZAR, M. AND GRETZINGER, S. 2005. *Costos y beneficios de la certificación forestal y mecanismos para la resolución de obstáculos comunes*. WWF Centroamérica and PROARCA. San José, Costa Rica. 80 pp.
- SCRIVEN, M. 1991. *Evaluation Thesaurus* (4th ed.). Newbury Park, CA: Sage. 391 pp.
- STEERING COMMITTEE OF THE STATE-OF-KNOWLEDGE ASSESSMENT OF STANDARDS AND CERTIFICATION (SCR). 2012. *Toward sustainability: the roles and limitations of certification*. RESOLVE, Inc., Washington, DC. 115 pp.
- STERN, E., STAME, N., MAYNE, J., FORSS, K., DAVIES, R., and BEFANI, B. 2012. *Broadening the range of designs and methods for impact evaluation*. Working Paper 38. Department for International Development (DFID), London. 127 pp.
- VAN KUIJK, M., PUTZ, F.E., and ZAGT, R. 2009. Effects of forest certification on biodiversity. Tropenbos International, Wageningen, the Netherlands, 94 p.
- VIDAL, N., and KOZAK, R. 2008. The recent evolution of corporate responsibility practices in the forestry sector. *International Forestry Review* **10**: 1–13.
- VINCENT, J.R. 2016. Impact evaluation of forest conservation programs: benefit-cost analysis, without the economics. *Environment and Resource Economics* **63**: 395–408. DOI 10.1007/s10640-015-9896-y
- VISSEREN-HAMAKERS, I.J., and PATTBURG, P. 2013. We can't see the forest for the trees—the environmental impact of global forest certification is unknown. *GAIA* **22**: 25–28.
- WINTERBOTTOM, R. 1990. *Taking Stock: the Tropical Forestry Action Plan after five years*. [Available at: <http://www.ponline.org/node/379246#sthash.6GagMHQs.dpuf>; accessed June 2016].