OECD Environment Working Papers No. 141

Evaluating the effectiveness
of policy instruments
for biodiversity: Impact
evaluation, cost-
effectiveness analysis and
other approaches

Katia Karousakis

<https://dx.doi.org/10.1787/ff87fd8d-en>

Evaluating the effectiveness of policy instruments for biodiversity: Impact evaluation, cost-effectiveness analysis and other approaches

An overview of methodologies and evidence across terrestrial and marine ecosystems

By Katia Karousakis, OECD

OECD Working Papers should not be reported as representing the official views of the OECD or of its member countries. The opinions expressed and arguments employed are those of the authors.

Authorised for publication by Rodolfo Lacy, Director, Environment Directorate.

Jel codes Q57, Q20,D61,D04

Keywords: Ecological Economics: Ecosystem Services • Biodiversity Conservation • Allocative Efficiency • Cost–Benefit Analysis, Microeconomic Policy: Formulation, Implementation, and Evaluation

OECD Environment Working Papers are available at
www.oecd.org/environment/workingpapers.htm.

OECD ENVIRONMENT WORKING PAPERS

OECD Working Papers should not be reported as representing the official views of the OECD or of its member countries. The opinions expressed and arguments employed are those of the author(s).

OECD Working Papers describe preliminary results or research in progress by the author(s) and are published to stimulate discussion on a broad range of issues on which the OECD works.

This series is designed to make available to a wider readership selected studies on environmental issues prepared for use within the OECD. Authorship is usually collective, but principal author(s) are named. The papers are generally available only in their original language -English or French- with a summary in the other language.

Comments on OECD Working Papers are welcomed, and may be sent to:

OECD Environment Directorate,
2, rue André Pascal, 75775 PARIS CEDEX 16, France
or by e-mail to env.contact@oecd.org

OECD Environment Working Papers are published on
www.oecd.org/environment/workingpapers.htm

This document and any map included herein are without prejudice to the status of or sovereignty over any territory, to the delimitation of international frontiers and boundaries and to the name of any territory, city or area.

© OECD (2018)

You can copy, download or print OECD content for your own use, and you can include excerpts from OECD publications, databases and multimedia products in your own documents, presentations, blogs, websites and teaching materials, provided that suitable acknowledgment of OECD as source and copyright owner is given. All requests for commercial use and translation rights should be submitted to rights@oecd.org

Acknowledgements

The report has been prepared by Katia Karousakis of the OECD Environment Directorate. Contributions from Will Symes on the inventory of studies are gratefully acknowledged. The author also thanks Dirk Rottgers, Daniel Nachtigall, Miguel Cardenas Rodriguez, Brilé Anderson and Rodney Boyd for useful comments on Section 2.1 and Annex 1. Suggestions on the paper from Antoine Dechezlepretre, Rob Youngman, Simon Buckle (all from the OECD Environment Directorate) and Hans Lundgren (OECD Development Co-operation Directorate) are also appreciated. The author would also like to thank delegates of the OECD Working Party on Biodiversity, Water and Ecosystems for helpful comments on earlier drafts of this paper. The author is responsible for any remaining omissions or errors.

Table of contents

Acknowledgements.....	3
Abstract.....	5
1. Evaluating the effectiveness of biodiversity policies.....	6
1.1. The need for evidence-based policy making.....	6
1.2. Impact evaluation and cost-effectiveness of biodiversity policies: where are we today?.....	8
1.3. Other approaches to evaluate the effectiveness of policy instruments for biodiversity.....	10
2. Methodologies for impact evaluation and cost-effectiveness analysis	12
2.1. Impact evaluation.....	12
2.2. Cost-effectiveness analysis	14
3. An inventory of biodiversity-relevant impact evaluation and cost-effectiveness studies.....	16
4. Policy insights and suggestions for further work	20
References	23
ANNEX 1: Overview of selected impact evaluation methodologies.....	29
ANNEX 2: Inventory of impact evaluation and cost-effectiveness analysis of biodiversity instruments.....	33

Figures

Figure 2.1. Conceptual diagram of impact evaluation.....	13
Figure 3.1. Number of impact evaluation studies by policy instrument.....	17
Figure 3.2. Number of impact evaluation studies by continent.....	18

Boxes

Box 1.1. Types of benchmarks against which to measure performance	11
Box 3.1. Findings from selected impact evaluation and CEA studies.....	19
Box 4.1. Integrating impact evaluation in the design and implementation of marine protected area monitoring	21

Abstract

This report provides an overview of methodologies to evaluate the effectiveness of policy instruments for biodiversity, covering impact evaluation, cost-effectiveness analysis and other more commonly used approaches. It then provides an inventory of biodiversity-relevant impact evaluation studies, across both terrestrial and marine ecosystems. The report concludes with lessons learned, policy insights and suggestions for further work.

JEL codes: Q57, Q20, D61, D04

Keywords: Ecosystem Services • Biodiversity Conservation • Allocative Efficiency • Cost-Effectiveness Analysis, Policy Design, Implementation, and Evaluation

Résumé

Ce rapport fournit une vue d'ensemble des méthodologies permettant d'évaluer l'efficacité des instruments de politique publique en faveur de la biodiversité, couvrant les études d'impact, les analyses coût-efficacité et d'autres approches plus couramment utilisées. Il présente ensuite un inventaire des études d'impact relatives à la diversité biologique des écosystèmes terrestres et marins. Le rapport se termine par des leçons tirées, recommandations politiques et des suggestions pour de futures recherches.

Codes JEL : Q57, Q20, D61, D04

Mots-clés : Services Écosystémiques • Conservation de la Biodiversité • Efficience allocative • Analyse coût-efficacité et évaluation de politiques publiques

1. Evaluating the effectiveness of biodiversity policies

1.1. The need for evidence-based policy making

When faced with a particular biodiversity challenge, environmental policy-makers must choose amongst alternative instruments for biodiversity conservation and sustainable use. A number of factors will affect this decision, including likely impact, cost-effectiveness, distributional issues, administrative ease, the ability to address uncertainties, and political feasibility (Goulder and Parry, 2008).

Evaluating the impact and cost-effectiveness of biodiversity policy instruments is crucial to ensure that the often limited resources available are spent so as to maximise outcomes. While the need to track biodiversity expenditure, identify funding gaps for biodiversity, and develop strategies to mobilise and scale-up finance is being increasingly recognised, much less attention has focussed on evaluating the impact of biodiversity policies (i.e., to determine what works, what doesn't – and why) and how this can be improved. This is true for both terrestrial and marine biodiversity instruments and finance and at various levels: domestic public finance, private finance, and biodiversity-related international development finance. Results from these kinds of rigorous analyses would allow governments, the private sector and donors to evaluate and adjust policies and future investments for improved biodiversity impact. Such studies are an important element of evidence-based policy making.

This need has also been recognised by the Parties to the Convention on Biological Diversity (CBD). Decision XIII/1 encourages Parties to undertake evaluations of the effectiveness of measures undertaken to implement the Strategic Plan for Biodiversity 2011-2020, to document experiences, including the methodologies applied, to identify lessons learned, and to provide this information to the Executive Secretary. To this end, the CBD Secretariat circulated a note on “Tools to evaluate the effectiveness of policy instruments for the implementation of the Strategic Plan for Biodiversity” [CBD/SBSTTA/21/7]. Recommendation XXI/6 on this issue was subsequently adopted by the CBD SBSTTA in December 2017, and a Decision adopted at CBD COP14 in November 2018.¹

This report builds on and responds to the CBD Secretariat note and the CBD COP14 Decision in three ways: 1) it provides a more in-depth overview of some of the more rigorous methodologies to evaluate the impact and cost-effectiveness of biodiversity policies, and discusses the use of other, more commonly used approaches to assess policy effectiveness; 2) it develops an inventory of biodiversity-relevant studies that have used these more rigorous approaches; and 3) it presents insights and lessons learned from the analysis and concludes with recommendations for further work.

Annual financial flows to biodiversity have been estimated at USD 51.8 billion in 2010 (Parker et al., 2012)², though these numbers are likely to have risen since then. Though not

¹ CBD SBSTTA Recommendation XXI/6 requests, *inter alia*, international organisations to share information on the methodologies used in evaluations of the effectiveness of measures taken to implement the Convention, including case studies, as well as lessons learned from these evaluations.

² Of this, general government budget was estimated at USD 25.6 billion; positive agricultural subsidies (from US and EU) at USD 7.8 billion; greening commodities (agriculture and fisheries

directly comparable with all the categories in Parker et al., more recent data indicates that potentially environmentally beneficial agriculture support in OECD countries was USD 13.7 billion in 2015³; bilateral biodiversity-related ODA was USD 8.3 billion on average per year in 2015-16⁴; and that revenue generated from biodiversity-relevant taxes amounted to approximately USD 7.4 billion in OECD countries in 2018⁵. Another estimate put total private capital committed to conservation investments at USD 4.7 billion in 2015, of which half was committed to water quality and quantity, and the rest equally split between habitat conservation, and sustainable food and fibre (Hammrick, 2016).

Looking across these various biodiversity policy instruments and finance flows, very little information is available on their actual *impact* in achieving biodiversity conservation and sustainable use, and which instruments tend to be more cost-effective. In line with Baylis et al. (2016), impact evaluation goes beyond monitoring programme inputs, outputs or indicators over time. It measures the causal effect of a specific policy, programme or intervention vis-à-vis a credible counterfactual scenario and seeks to understand the conditions under which this effect arises (Ferraro and Hanauer, 2014). Impact evaluation focuses on disentangling the effects attributable to a particular policy intervention (e.g. protected areas) on an outcome variable (e.g. deforestation) from broader changes in a region (e.g. widespread development or government policies) (Ahmadia, 2015). More generally, according to one source, the resources currently allocated to monitoring and evaluation — on average, less than 5% of a conservation project's budget — do not come close to what is needed to satisfy the increasing demand from policymakers for more and higher-quality evidence on the impacts of conservation and development interventions (McKinnon et al., 2015a)⁶. This need for better evidence-based policy making holds true in both the terrestrial and marine biodiversity context.

In the context of development co-operation, the OECD Development Assistance Committee (DAC) defines impact as “the positive and negative changes produced by a development intervention, directly or indirectly, intended or unintended. This involves the main impacts and effects resulting from the activity on the local social, economic, environmental and other development indicators.” (OECD, 2010)⁷. Other definitions of impact evaluation also exist such as that of the World Bank Development Impact Evaluation (DIME) initiative, which states: impact evaluations compare the outcomes of a programme against a counterfactual that shows what would have happened to beneficiaries without the programme.

certification) at USD 6.6 billion; bilateral biodiversity-related ODA at USD 6.3 billion; biodiversity offsets at USD 3.2 billion; and direct biodiversity (i.e. user) fees at USD 0.4 billion.

3 OECD Producer Support Estimate (PSE) database, based on data as of November 2016.

4 OECD DAC Creditor Reporting System (CRS) database, as of July 2018.

5 OECD Policy Instruments for the Environment (PINE) database, based on data as of October 2018. www.oecd.org/pine. See also: [Tracking Economic Instruments and Finance for Biodiversity](#).

6 Monitoring and evaluation is distinct from, but can make valuable contributions to, impact evaluation (see Perrin, 2012 for a discussion).

7 The five principles of the OECD DAC Evaluation Criteria for development assistance consist of: relevance, effectiveness, efficiency, impact and sustainability. The second focus of this report on cost-effectiveness relates most closely to the DAC criterion on efficiency, which measures the outputs – qualitative and quantitative – in relation to the inputs.

The need for better impact evaluation and cost-effectiveness analysis is also implicit in various national and supra-national policy frameworks, such as EU Marine Strategy Framework Directive (MSFD) and Australia's biodiversity strategy. The MSFD, for example, explicitly requires the member states to assess the present state of the sea in question and develop a national Programme of Measures (PoM) designed to narrow and eventually close the gap between the current and desired state of the sea. The member states must show that the chosen PoM is cost-effective (Oinonon, 2016). This approach is also consistent with the broad goal of Australia's Biodiversity Conservation Strategy 2010–2030 to deliver conservation initiatives in a cost-effective manner. Moreover, a number of development agencies are supporting the use of impact evaluation to assess the effectiveness of interventions abroad. Examples include the Japan International Cooperation Agency which has supported impact evaluation studies to evaluate forest conservation programmes in Ethiopia (see Takashi and Todo, 2013) and the Global Environmental Facility which has undertaken an impact evaluation of GEF support to protected areas and protected area systems (GEF, 2016). The German Institute of Development Evaluation (DEval) was created in 2012 to carry out independent evaluations of Germany's development interventions, providing the foundations for informed policy making⁸.

1.2. Impact evaluation and cost-effectiveness of biodiversity policies: where are we today?

More than 10 years ago, Ferraro and Pattanayak (2006) made a call for better empirical evaluation of biodiversity conservation investments. They argued that evaluation focus must shift from “inputs” (e.g., investment dollars) and “outputs” (e.g., training) to “outcomes” produced directly because of conservation investments (e.g., species and habitats), stating that “*In the field of program evaluation, one lesson is paramount: you cannot overcome poor quality with greater quantity.*” A subsequent review conducted in 2012 re-confirmed the claim that credible evaluations of common biodiversity instruments continue to be rare (Miteva, Pattanayak and Ferraro, 2012)⁹. Calls for more effective conservation-related development finance have also been made (e.g. Waldron et al. 2013; Richertzhagen et al. 2016).

Albeit still sparse, additional rigorous studies are emerging however (see section 3, and Annex 2 for an inventory). Law (2016) indicates that today many funders now request or encourage rigorous conservation evaluation as a condition of funding, and conceptual and technical how-to literature is increasingly available (Ferraro 2012, World Bank Group 2013, Fisher et al. 2014). There is also an expanding drive and capacity to collate the required data (Ferraro and Pressey 2015; Bare et al. 2015). The lack of studies using rigorous impact evaluation methodologies may be partly explained by the fact that their use was not planned for in the design of the policy, meaning there is no counterfactual to measure against. Institutions, including governments and NGOs, are now increasingly

⁸ See for example: Leppert, Hohfeld, Lech and Wencker (2018), *Impact, Diffusion and Scaling-Up of a Comprehensive Land-Use Planning Approach in the Philippines. From Development Cooperation to National Policies*, German Institute for Development Evaluation (DEval), Bonn.

⁹ To date, these methods have more commonly been applied in the field of development economics, public health, and education.

recognising the need for rigorous evaluations within the context of evidence based policy, even if they uncover negligible or negative impacts (McKinnon et al. 2015b).

Other initiatives that consolidate relevant information and aim to make it more accessible to policy-makers are also emerging. For example, the Initiative for Impact Evaluation (www.3ieimpact.org) has collected more than 4 260 impact evaluation studies of which 240 are environmentally-relevant. The Poverty Action Lab (J-PAL) initiative documents 843 impact evaluation (specifically randomised) studies, about 60 of which are relevant to the environment and energy (www.povertyactionlab.org/evaluations). The World Bank Development Impact Evaluation (DIME) project, whose objective is to increase the use of impact evaluation in the design and implementation of public policy, also collects this information (<http://www.worldbank.org/en/research/dime>). The Network of Networks on Impact Evaluation (NONIE), comprised of the OECD DAC Evaluation Network, the United Nations Evaluation Group (UNEG), the Evaluation Cooperation Group (ECG), and the International Organisation for Cooperation in Evaluation (IOCE), was formed in 2006 to promote quality impact evaluations. The group has also developed NONIE guidance on impact evaluation (Leeuw and Vaessen, 2009).

Systematic reviews in this area are also being developed. Systematic reviews examine the results of a synthesis of individual impact evaluation studies to identify general relationships and treatment effects. The Collaboration for Environmental Evidence (www.environmentalevidence.org) has published systematic reviews of the effect of protected areas, payment for ecosystem services and aspects of forest management on various human welfare, habitat and species preservation outcomes (Baylis et al. 2016; see Samii et al. 2015 for an example). Puri et al. (2016) have identified 8 systematic reviews in the context of forest conservation interventions. Including a broader variety of approaches to evaluate effectiveness, the OECD DAC Evaluation Resource Centre (DEReC) has collected over 3000 evaluation reports from its member's development agencies, including on environment, forestry and agriculture (www.oecd.org/derec).

Moving beyond quantitative impact evaluation to determine what works and what doesn't, a related field is that of theory-based impact evaluation. Theory-based evaluation aims to ascertain, through a mixed methods approach, the question of *why* a programme or policy did or did not work. It examines the assumptions underlying the causal chain from inputs to outputs, outcomes and impacts. According to White (2009), the six key principles of theory-based impact evaluation are:

- 1. Map out the causal chain (programme theory)
- 2. Understand context
- 3. Anticipate heterogeneity
- 4. Rigorous evaluation of impact using a credible counterfactual
- 5. Rigorous factual analysis
- 6. Use mixed methods¹⁰

¹⁰ According to Bamberger (2012), "Mixed methods evaluations seek to integrate social science disciplines with predominantly quantitative (QUANT) and predominantly qualitative (QUAL) approaches to theory, data collection, data analysis and interpretation. The purpose is to strengthen the reliability of data, validity of the findings and recommendations, and to broaden and deepen our understanding of the processes through which program outcomes and impacts are achieved, and how these are affected by the context within which the program is implemented."

While impact evaluation is important in order to identify what works and what doesn't, and can be combined with more qualitative methods of evaluation to determine why the policy or programme did or did not work, the next logical step is ex-post assessment of the cost-effectiveness of biodiversity conservation and sustainable use instruments. Having identified the impact of a given policy in meeting its intended objectives, information on the finance needed to achieve this would enable a cost-effectiveness analysis. Ferraro et al. (2013) point out the glaring lack of cost data in impact evaluation studies and Vincent (2016) discusses the challenges with this in the context of forest conservation where costs and benefits are spatially heterogeneous.

In a global review investigating the integration of economics in agri-environment scheme (AES) evaluation, Ansell et al. (2016) find that fewer than 15% of studies consider cost-effectiveness in scheme evaluation (i.e. 31 studies). Given that AES are the focus of significant investment around the world, with agri-environmental investment in many countries often equal to, or surpassing that of other conservation expenditure (Batáry et al., 2015), further work in this area is also important¹¹.

1.3. Other approaches to evaluate the effectiveness of policy instruments for biodiversity

The impact evaluation methods described above are *attributional*, that is, they support inferences about the causal relationship between the treatment and the indicators. Other more commonly used approaches to evaluate effectiveness are *non-attributional*, that is, not supportive of any causal claim but assessing the level of the indicators against other benchmarks (Coglianese, 2012). Non-attributional evaluations are often used in performance measurement, strategic management, and budgeting practices (U.S. OMB, 2010, p. 83). They typically compare current measurements of performance with one or more of the benchmarks highlighted in Box 1.1 (Coglianese, 2012).

¹¹ Moreover, a review of impact evaluations in the agriculture sector conducted by IEG (2011) identified 86 impact evaluation studies, 50% of which used quasi-experimental or RCT methods. Of the total, 14 percent examined natural resource management relevant interventions, defined as those that sought to improve farmer knowledge and adoption of new technologies and conservation techniques. They included soil and water conservation (9 percent), systems of crop management (4 percent), and integrated aquaculture-agriculture (1 percent).

Box 1.1. Types of benchmarks against which to measure performance

- *Treatment goals.* Do the indicators show levels that meet regulatory officials' goals or targets (e.g., have nutrient pollution levels decreased to the level desired), regardless of whether caused by the regulation?
- *"Acceptable" levels.* Do the indicators show that the problem has been reduced sufficiently, such as to below a morally tolerable threshold that has been independently determined (e.g., reducing nutrient pollution to below a safe level)?
- *Historical benchmarks.* Are the indicators better today than they were before (regardless of whether the treatment actually caused any of the change)?
- *Other jurisdictions.* Are the indicators in the jurisdiction with the regulation different than in other jurisdictions (again, regardless of whether the regulation contributed to any of the difference)?

Source: Coglianese (2012), "Evaluating the impact of regulation and regulatory policy", Expert Paper No.1, OECD Paris.

The ideal indicators to use in evaluating a policy will always be measures of impacts, as these are what matter in the end. However, this does not mean there is not a role, nor even a significant one, for indicators on inputs, processes, and other outcomes.

The OECD Framework for Regulatory Evaluation describes how different types of indicators can be used to create an overarching method of evaluating policy performance, ranging from inputs, process, outputs, outcomes and impacts (OECD, 2014). This is in line with suggested approaches to e.g. monitor progress towards biodiversity mainstreaming (see OECD, 2018). As part of the OECD (2014) work, a survey was conducted to obtain information on OECD country practices with respect to regulatory policy evaluation. The report states:

"While it is widely agreed that systematic evaluation of existing regulations through ex post impact analysis is necessary to ensure that regulations are effective and efficient, few countries are actually performing it on a systematic basis. In particular, few countries assess whether underlying policy goals of regulation have been achieved, whether any unintended consequences have occurred and whether there is a more efficient solution to achieve the same objective. A more frequent practice in OECD countries is partial ex post assessment focusing exclusively on regulatory burdens".

Nevertheless they conclude that the application of the Framework is within reach for all OECD countries. While not all countries will be in a position to collect information at all stages, there is already great value in collecting data on the actual implementation of regulatory policy (outputs) and its impact on regulations (intermediate outcomes). The latest insights on regulatory policy evaluation are available in the 2018 Regulatory Policy Outlook (OECD, 2018).

Also of relevance in this regard is the International Organisation of Supreme Audit Institutions (INTOSAI) under which the Working Group on Environmental Auditing collects and makes publicly available reports on environmental audits. Many of these are relevant to terrestrial and marine biodiversity (see <https://www.environmental-auditing.org/>).

2. Methodologies for impact evaluation and cost-effectiveness analysis

2.1. Impact evaluation

Impact evaluation asks what is the causal effect of a programme P on an outcome Y for unit i ? Causal impacts are unit specific, where i can refer to locations, firms, households, individuals, species, etc (Ferraro and Hanauer, 2014). This can be illustrated by a formula that the causal impact (Δ) of a programme (P) on an outcome (Y) for unit i is the difference between the outcome (Y) *with* the programme (in other words, when $P = 1$) and the same outcome (Y) *without* the programme (that is, when $P = 0$).

$$\Delta_i = (Y_i | P = 1) - (Y_i | P = 0).$$

The key challenge is to determine what this counterfactual is i.e., what would the alternative state for unit i have been without the programme. The difference between the two states cannot be directly observed but can be inferred based on empirical evidence and some assumptions. It involves moving from the unit level to the group level, and identifying a treatment group and a control group (i.e. a valid counterfactual) that are statistically similar, on average, or at least follow the same trend, in the absence of the programme. Two common problems that arise in the construction of a valid counterfactual are:

Omitted variables: the risk of attributing to the programme or policy effects due to other intervening factors external to the programme/policy (such as the change in economic conditions).

Selection bias: the risk of attributing to the programme/policy the effect of non-observable variables that affect participation to the programme/policy through the selection mechanism (ability, motivation, environmental awareness, etc.).

Figure 2.1. Conceptual diagram of impact evaluation

Comparison of the situation actually observed and the counterfactual situation

Source: JICA operation evaluations system.

http://www.jica.go.jp/english/our_work/evaluation/reports/2014/c8h0vm00009ga6st-att/part1_2014.pdf

Two common empirical designs employed by natural scientists to assess the performance of biodiversity interventions rely on comparisons of outcomes in areas (a) with and without exposure to a biodiversity policy instrument, or (b) before and after a biodiversity policy instrument is implemented. ‘With–without’ analyses implicitly assume that (i) the areas with and without the biodiversity policy are similar in terms of their expected outcomes in the absence of the biodiversity policy (i.e. similar in characteristics that affect outcomes, such as accessibility, suitability for agriculture, and proximity to markets) and (ii) there are no spillover effects from the biodiversity policy to ‘unexposed’ areas. ‘Before–after’ analyses assume that the outcome level (or its trend) before a policy is enacted would remain constant after the policy is enacted (Nagendra, 2008) and that there is no selection bias in targeting the policy (Miteva et al., 2012).

Various methods have evolved to rigorously evaluate the causal impact of policies or programmes (referred to as programme impact evaluation in the literature). The methods are classified as experimental and quasi-experimental designs¹². Experimental designs construct a control group through random assignment. In contrast quasi-experimental designs construct a comparison group via methods such as matching, instrumental variables, and difference-in-difference (Pattanayak, 2009; Gertler et al. 2016). These four

¹² Some use a broader definition of impact evaluation and also include non-experimental designs. These look systematically at whether the evidence is consistent with what would be expected if the intervention was producing the impacts, and also whether other factors could provide an alternative explanation. See e.g., http://www.betterevaluation.org/en/themes/impact_evaluation.

methods are discussed in further detail in Annex 1 in order to illustrate the types of issues that can arise and how they can be addressed. Other methods are also available¹³ and several guidance documents and handbooks exist. The interested reader can refer to Gertler et al. (2016), Ferraro and Hanauer (2014), Lance et al. (2014), and Leeuw and Vaessen (2009).

2.2. Cost-effectiveness analysis

It is critically important that impact evaluation be complemented with information on the cost of the project, programme or policy being evaluated. Once impact evaluation results are available, they can be combined with information on programme costs to analyse cost-effectiveness (if possible, a fully-fledged cost-benefit analysis should be carried out (OECD, 2018; Gertler et al., 2016).

A cost-effectiveness analysis (CEA) seeks to find the alternative activity, process or intervention that minimises resource use to achieve a desired result. An ex-post CEA addresses the question of how far objectives have been achieved, and at what cost. The cost-effectiveness of a policy is calculated by dividing the annualised costs of the option by a quantified measure of the physical effect, such as animal or plant species recovered, tonnes of emissions of a given pollutant reduced, kilometres of river length restored, and so on (Gorlach, n.d.). Given that the physical effect of a biodiversity intervention can be measured (or deduced from) by a variety of metrics, it is therefore necessary to limit the comparison of alternatives to those that have similar goals.¹⁴

There is some discussion on which types of costs should be considered in a CEA, ranging from the purely financial private costs (investment and operational costs) of specific measures to general equilibrium-estimates of costs to the wider economy, including efficiency losses (foregone welfare). If projects or interventions are small however, they are less likely to have impacts on the wider economy and the latter may not be needed. Generally, the types of costs should include start-up costs, programme costs, capital costs, personnel costs and opportunity costs. All the costs should be discounted to their present value.

An issue to consider is whether measures of effectiveness should be discounted even though they are in non-monetary terms.¹⁵ Regarding the temporal dimension of effectiveness, there is no guidance on whether some type of discounting should be applied as well.

Overall, McVittie et al. (2014) find that CEA is more routinely applied in health economics (Gold et al., 1996), and has so far only been applied to the evaluation of conservation programmes on a handful of occasions (e.g. Laycock et al., 2009, 2011, 2012; Montgomery et al., 1994; Macmillan et al., 1998; Fairburn et al., 2004).

¹³ For example, regression discontinuity design.

¹⁴ In the field of health economics, for example, the metric developed for health assessment includes the Quality Adjusted Life Year (QALY) and the Value of Statistical Life (VSL), which collapses many otherwise incommensurate health states to a common metric or scoring system.

¹⁵ Canada uses CEA in the application of its Species At Risk Act (SARA) Action Plans where the benefits are in non-monetary terms. These are expressed in terms of probability, and are therefore not discounted. The approach is based on prioritising threat management for biodiversity, following Carwardine et al. (2012).

A study on community-based forest management in India by Somanathan et al. (2009) provides an example of combining impact evaluation results with information on conservation costs, as does the study on protected areas and PES in Mexico by Sims and Alix-Garcia (2016).

There are also seemingly less guides or manuals available on how to undertake a CEA than there are for IE, with some exceptions including Dhaliwal et al. (2011) with applications for education¹⁶. As indicated above, the main challenge in CEA is that investment comparisons and choices need to be based on a constant outcome metric. In other words, species and landscape scale outcomes need to be scaled using a comparable scale. Another challenge is that CEA generally only looks at one outcome and does not consider other impacts such as co-benefits.

¹⁶ Guidance documents on cost-benefit analysis do exist (see e.g. Australian Government, 2016).

3. An inventory of biodiversity-relevant impact evaluation and cost-effectiveness studies

A review of the literature yielded more than 80 impact evaluation and CEA studies that are relevant to biodiversity conservation and sustainable use (see Annex 2). Though there are likely to be additional impact evaluation and CEA studies available, this sample is nevertheless likely to represent the majority of studies that exist at this time. The policy instruments that have been evaluated via these methods are:

- Protected Areas – terrestrial and marine
- Payments for Ecosystem Services (including REDD+)
- Agri-environment schemes
- Eco-certification (e.g. forests and coffee)
- Community-based forest management
- Integrated conservation and development projects (ICDP)
- Zoning policy
- Moratorium (on oil palm, timber and logging concessions)
- Bans on logging
- Black-listing
- Social marketing campaigns
- Law enforcement
- Other forms of technical assistance

Using SCOPUS (the largest database of peer-reviewed literature) to search through the literature, no impact evaluation studies were found in the context of biodiversity offsets, biodiversity-relevant taxes, tradable permits, and cross-compliance.

For each of the studies, the following information is collected (Annex 2):

- Policy instrument
- Location
- Scope (e.g. national, sub-national, regional)
- Methodology (e.g., matching) and/or CEA
- Unit of measurement
- Impact
- Cost (i.e., included or not included)
- Reference

Looking across the existing studies identified in this review, only three studies (of more than 80) focus on the marine environment. These are Ahmadia et al. (2015) on marine protected areas¹⁷; Miteva et al. (2015) on the effectiveness of marine protected areas and species management PAs on conserving mangroves and reducing blue carbon emissions

17 Ahmadia et al (2015) collected data to conduct the baseline analysis for an impact evaluation. Time-series data is planned to be collected in order to undertake a full impact evaluation study.

(via carbon sequestration); and Verissimo et al. (2018) on social marketing campaigns targeted to fishers in MPAs. All other studies identified focus on terrestrial ecosystems.

Descriptive statistics on the policy instruments and the geographic representation are presented below. With regard to the former, the majority of IE studies focus on protected areas and on PES (Figure 3.1). In terms of geographical representation, most biodiversity-relevant impact evaluation studies have been conducted in developing countries notably in Latin America, followed by S.E. Asia (Figure 3.2).

Figure 3.1. Number of impact evaluation studies by policy instrument

Notes: Of the 36 impact evaluation studies on protected areas, 2 were on marine protected areas. PES = Payments for Ecosystem Services. AES = Agri-environment schemes. CBFM = Community-based forest management. REDD+ = Reducing emissions from deforestation and forest degradation. ICDP = integrated conservation and development projects.

Figure 3.2. Number of impact evaluation studies by continent

Turning to the measurement units that have been used to evaluate impact, most of the studies examine changes in forest cover or in deforestation rates. This is true predominantly for studies evaluating community-based forest management schemes, (terrestrial) protected areas, payments for ecosystem services programmes, and ICDP. The agri-environment and farmland conservation schemes evaluate either farmers' behaviour, or hectares of cover crops planted. Very few of the studies use more specific impact indicators. Exceptions include an index of biodiversity decline (Waldron et al. 2017), a biodiversity index value (Hily et al. 2015); the percentage of nests successfully fledgling (Clements et al. 2013); a change in population index (Saldarriaga-Isaza, 2007), species recovery rates (Ferraro, McIntosh, and Ospina, 2007), and changes in fish biomass and in marine habitat quality (Ahmadia et al. 2015).

The results of a few of these studies are summarised in Box 3.1 below. An overview of all other results is provided in Annex 2.

Looking across the results of the effectiveness of PES and protected areas (where a larger number of studies are available), nearly all of the studies evaluating PES programmes indicate that such payments have resulted in a positive environmental impact. While this is interesting in-and-of itself, it would also have been interesting if the studies had provided information on the average payment/unit measured (e.g. hectare or type of ecosystem service) for each of the PES schemes, and if possible, information on the opportunity costs of land.

Turning to protected areas, the results summarised in Annex 2 suggest that these have had a positive impact, but the results are perhaps less conclusive than those of PES. To some extent this also raises the question of associated monitoring and enforcement efforts across protected areas. It would be interesting, for example, to better understand if the results depend on the average PA spending per hectare. Overall, information on payments (or other form of funding) would be interesting as, for example, Ferraro, McIntosh, and Ospina (2007) used matching methods to analyse the impact of the U.S. Endangered Species Act on species recovery rates and found significant improvements in recovery rates but only

when the listing was combined with substantial government funding for habitat protection. Nelson et al. (2017) evaluate the impacts of critical habitat designation under the U.S Endangered Species Act on land cover change and find that, on average, it does not have impact. They acknowledge that their study does not take into account any land management and water use impacts of designation.

Box 3.1. Findings from selected impact evaluation and CEA studies

In a comprehensive national-scale impact evaluation and comparison of **protected area parks** versus **payments for ecosystem services** in Mexico, Sims and Alix-Garcia (2016) find that both policies conserved forest, generating an approximately 20-25% reduction in expected forest cover loss. Looking at all protected area types and the federal PES programme, they find all policies generated avoided deforestation from a mix of low and high cost land. They also found that there is no clear policy winner with respect to **cost-effectiveness**.

Using propensity score matching and Difference-in-Difference, Andam et al. (2008) evaluate the impact on deforestation of Costa Rica's Protected Area system between 1960 and 1997. They found that protection reduced deforestation: approximately 10% of the PAs would have been deforested had they not been protected. In contrast, the conventional approaches to evaluation overestimated avoided deforestation by over 65%.

In a study examining the environmental impacts of forest certification in Indonesia, Miteva, Loucks and Pattanayak (2015) find that between 2000 and 2008, Forest Stewardship Council (FSC) certification reduced aggregate deforestation by 5 percentage points. Three other studies of certification schemes also found positive impact, whereas one (Villalobos et al. 2018) did not.

Bare et al. (2015) assess the impact of **international conservation aid** on deforestation in sub-Saharan Africa and find that conservation aid is associated with higher rates of forest loss after one- or two-year lags. A similar result holds for Protected Area extent, suggesting possible displacement of deforestation from PAs (i.e. leakage).

A cost-effectiveness of Natura 2000 contracts in forests has been undertaken by Hily et al (2015), and of a **protected area** in Chile by Saldarriaga-Isaza et al (2010). Laycock et al (2009) estimated the cost-effectiveness of the **UK Biodiversity Action Plan**. The present value cost of implementing the 44 Species Action Plans for which cost data were available over the 10-year period in question varied from GBP 500 to GBP 7,000,000. The analysis enabled to make recommendations regarding how the finance could better be allocated to reach objectives at lower total cost. Earlier studies include a cost-effectiveness analysis of an **endangered species management** in New Zealand (Fairburn et al., 2004), and of **woodland ecosystem restoration** (Macmillan, 1998).

Based on a review of the literature, a similar analysis in the context of fisheries does not seem to exist. In Denmark, Kronbak and Vestergaard (2013) apply environmental cost-effectiveness analysis to a selective **fishing gear policy** case in Danish mixed trawl fisheries in Kattegat and Skagerrak. They compare the outcomes of two different gear types (referred to as 90/120 mm vs. grid) to the baseline. Cost data is not included however.

4. Policy insights and suggestions for further work

Rigorous impact evaluation studies offer robust evidence on whether a policy intervention has had the desired impact on biodiversity. However, such studies are not needed for all interventions. A practical consideration in the use of impact evaluation is cost. Impact evaluations are expensive and it is neither required nor possible to undertake impact evaluations of all programmes or policies. The use of impact evaluation has, however, garnered substantially greater attention in the field of development economics, health and education (Sabet and Brown, 2018), and is being used to inform and adjust policy-making in these fields (World Bank, 2009). Moreover while the costs of an impact evaluation can be high, designing or adjusting programmes without information on whether or not they work can be significantly more costly. Tax-payers also demand accountability for results. There is a need therefore to build a strong evidence base across all sectors - including those relevant to biodiversity - in a variety of contexts to provide guidance for policy-makers (OECD, 2006). The general lack of impact evaluation studies in the field of environment, and more specifically biodiversity, implies that these fields are not keeping up with best practice. There is a risk, therefore, of wasting money; not optimising biodiversity conservation and sustainable use spending; and not adjusting biodiversity programmes to achieve greater impacts on the ground.

The analysis in this report finds that most of the biodiversity-relevant impact evaluation studies are being conducted in developing countries. This is likely to be due to the need for development co-operation providers to show accountability for the financial and other resources invested, and the fact that impact evaluation is more common overall in the development field. Most of the existing biodiversity-relevant impact evaluation studies focus on protected areas and payments for ecosystem services. Very few impact evaluation studies exist however on other policy instruments that are also commonly used across countries. Given the large budgetary investments, for example, of agri-environmental schemes, it is perhaps surprising that few if any impact evaluation studies exist that aim to examine their impact on the environment. Moreover, nearly all of the studies identified focus on terrestrial ecosystems, with only a very small minority of studies on marine ecosystems (i.e. three studies, namely: Ahmadi et al (2015) and Miteva et al (2015) on impacts of MPAs; and Verissimo et al (2018) on the impact of social marketing campaigns in fisheries). Based on the analysis in this paper, the following recommendations are made for further work in this area:

Governments could endeavour to develop a strategic approach to scale up impact evaluation and cost-effectiveness analysis so as to build a stronger evidence base for more environmentally- and cost-effective biodiversity policy instruments. This could include considerations of geographic representability, ensuring a good balance between different policy instruments and of different types of terrestrial and marine ecosystems, and ideally prioritising larger initiatives that are being planned¹⁸.

18 As an example, criteria used by DEval to determine whether to undertake an impact evaluation study include: coverage; innovation and learning potential; risk; and strategic importance. For more information see http://www.deval.org/en/evaluations.html?file=files/content/Dateien/Evaluierung/DEval_Evaluierungsprogramm_2018_EN_final.pdf.

In terms of ways forward, Puri et al. (2016) state that targeted investment in baseline data (qualitative and quantitative) on intended outcomes of conservation interventions (ecological, social and economic) may substantially increase the ability to evaluate interventions in geographic areas that lack substantial secondary data archives. Greater efforts are therefore needed to systematically integrate impact evaluation and cost-effectiveness analysis in the design and implementation of biodiversity policy monitoring. More generally, Gertler et al. (2016) call for better selection of outcome and performance indicators, where good indicators are SMART (specific, measurable, attributable, realistic, and targeted). Le Velly and Dutilly (2016) propose guidelines and discuss methodological choices in designing impact evaluation for PES programmes and Ahmadia (2015) discusses how impact evaluation monitoring can be integrated into the design of Marine Protected Areas (see Box 4.1). In another paper, Peterson et al. (2018) provide a quantitative framework for evaluating the impact of biodiversity offset policies.

Box 4.1. Integrating impact evaluation in the design and implementation of marine protected area monitoring

In 2011, scholars and practitioners recognised the potential to modify the ecological monitoring programme implemented in the Bird's Head Seascape (BHS) in Indonesia to enable quasi-experimental causal inference at the Seascape scale. Considerable monitoring efforts took place from 2009 to 2014 inside both no-take and use zones of the six Marine Protected Areas (MPAs), and in 2012, in areas outside of MPAs, to document baseline ecological conditions (fish and benthic attributes) of coral reefs.

Ecological indicators were selected to reflect management goals, inform policy-makers, and be useful as indicators of ecosystem health and fish populations. The following indicators were selected for inclusion: (i) overall biomass of key fisheries species (Lutjanidae, Haemulidae, Serranidae), (ii) biomass of herbivorous fish species (Acanthuridae, Scaridae, Siganidae), and (iii) habitat quality (ratio of hard coral cover to rubble and algae cover).

Matching methods were then used: MPAs in the BHS have been strategically designed and (non-randomly) placed. To avoid observable selection bias, a quasi-experimental design was adopted, and a tiered matching approach was applied (coarse matching followed by statistical matching) to identify comparable control sites, which allowed the estimation of the counterfactual (i.e. changes in fish biomass that would have occurred if no MPA were established).

Source: Ahmadia, G., et al. (2015), Integrating impact evaluation in the design and implementation of monitoring marine protected areas, *Phil. Trans. R. Soc.*, B 370.

Given the multidimensionality of biodiversity (covering both terrestrial and marine ecosystems) and ecosystem services, the development of further, more comprehensive guidance in this area is likely to be useful. Such a guidance document could, for example, highlight the various types of impact indicators that could possibly be examined, organised by ecosystem types, and when the data needs to be collected. Biodiversity-relevant policy practitioners designing new policies or programmes could then use this as a check-list of issues to consider, so as to enable the establishment of a counter-factual, and eventually,

the undertaking of an impact evaluation study. Overall, much greater attention is being paid to these methods also in academic circles, and the number of impact evaluation studies covering all fields continues to rise (Sabet and Brown, 2018). Government policy-makers and practitioners interested in ensuring that actions and finance invested is achieving real impact on the ground could therefore also enable academics to undertake such rigorous studies if relevant data is collected. Fostering stronger collaboration in the science-policy interface (i.e. between academia and government) may also help to move from an ad-hoc emergence of biodiversity-relevant impact evaluation and cost-effectiveness analysis studies, to a more strategic one.

References

- Ahmadia, G., et al. (2015), Integrating impact evaluation in the design and implementation of monitoring marine protected areas. *Phil. Trans. R. Soc. B* 370.
- Ansell, D., et al. (2016), The cost-effectiveness of agri-environment schemes for biodiversity conservation: A quantitative review. *Agriculture, Ecosystems and Environment*, 225, 184-191.
- Arata, L., Sckokai, P. (2016), The impact of agri-environmental schemes on farm performance in five E.U. member states: a did-matching approach. *Land Econ.* 92, 167–186.
- Arriagada, RA., et al. (2012), Do Payments for Environmental Services Affect Forest Cover? A Farm-Level Evaluation from Costa Rica, *Land Economics*, 88 (2).
- Australian Government (2016), *Guidance Note: Cost Benefit Analysis*.
www.pmc.gov.au/sites/default/files/publications/006-Cost-benefit-analysis.pdf.
- Bamberger, M. (2012), *Introduction to Mixed Methods in Impact Evaluation*. Impact Evaluation Notes, No. 3. InterAction.
- Bare, M. (2015), Assessing the impact of international conservation aid on deforestation in sub-Saharan Africa. *Environmental Research Letters*, 10 (12).
- Barral, M., et al. (2015), Quantifying the impacts of ecological restoration on biodiversity and ecosystem services in agroecosystems: A global meta-analysis. *Agriculture, Ecosystems and Environment*, 202: 223–231.
- Baylis, K., et al. (2016), Mainstreaming impact evaluation in nature conservation. *Conservation Letters*, 9 (1).
- Blackman, A. (2012), Ex Post Evaluation of Forest Conservation Policies Using Remote Sensing Data. *Resources for the Future*, Discussion Paper 12–13. Washington D.C.
- Cawardine, J., et al (2012), Prioritising Threat Management for Biodiversity Conservation, *Conservation Letters*, 5 (3).
- Coglianesi, C. (2012), “Measuring Regulatory Performance: Evaluating the impact of regulation and regulatory policy”. Expert Paper No.1, OECD, Paris, www.oecd.org/gov/regulatory-policy/1_coglianesi%20web.pdf.
- Costedoat, S., et al.. (2015), How Effective Are Biodiversity Conservation Payments in Mexico? *PLoS ONE*, 10(3).
- Cropper, M., Puri, J., Griffiths, C. (2001), How the location of roads and protected areas affects deforestation in North Thailand. *Land Economics*, 77.
- Dhaliwal, I., et al. (2011), *Comparative cost-effectiveness analysis to inform policy in developing countries: a general framework with applications for education*. MIT Press, Cambridge, MA.
- De Lima, ACB., Luiz A., Keppe N., Alves M.C. (2008), Impact of FSC forest certification on agro-extractive communities of the state of Acre, Brazil.
- Doole, G., et al. (2014), Determinants of cost-effectiveness in tender and offset programmes for Australian biodiversity conservation. *Land Use Policy*, 36, 23-32.
- Fairburn, G., Hughey, K., Cullen, R., (2004), Cost effectiveness of endangered species management: the kokako (*Callaeas cinerea*) in New Zealand. *New Zealand Journal of Ecology*, 28, pp. 83–91.

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. London*, B 370.

Ferraro, P. J. and Hanauer, M., (2014), Advances in measuring the environmental and social impacts of environmental programs. *Ann. Rev Environ. Resources*, 39 495–517.

Ferraro, P.J., Hanauer, M., Miteva, D.A. et al. (2013), More strictly protected areas are not necessarily more protective: evidence from Bolivia, Costa Rica, Indonesia, and Thailand. *Environ Res Letters*, 8 (2).

Ferraro, P.J., McIntosh, C., Ospina, M., (2007), The effectiveness of the US endangered species act: an econometric analysis using matching methods. *Journal of Environmental Economics and Management*. 54 (3), 245–261.

Ferraro, P.J., Pattanayak, S.K. (2006), Money for nothing? A call for empirical evaluation of biodiversity conservation investments. *PLoS Biol*, 4(4):e105.

Ferraro, P.J. (2012), Experimental project designs in the global environment facility: Designing projects to create evidence and catalyze investments to secure global environmental benefits. A STAP Advisory Document. Global Environmental Facility, Washington, DC.

Fisher, B., et al. (2015), Cost-effective conservation: calculating biodiversity and logging trade-offs in Southeast Asia. *Conservation Letters*, 4 (6).

Global Environment Facility Independent Evaluation Office (GEF IEO) (2016), *Impact Evaluation of GEF Support to Protected Areas and Protected Area Systems*, Evaluation Report No. 104, Washington D.C.

Gertler, P., et al. (2016), *Impact Evaluation in Practice, Second Edition*. Inter-American Development Bank and World Bank. Washington, DC: <https://openknowledge.worldbank.org/handle/10986/25030> License: CC BY 3.0 IGO.

Goulder, L., and Parry, I. (2008), Instrument Choice in Environment Policy. *Rev Environ Econ Policy*, 2 (2): 152–174.

Hamrick, K. (2016), *State of Private Investment in Conservation 2016: A Landscape Assessment of an Emerging Market*. Ecosystem Marketplace.

Hily, E., et al. (2015), Assessing the cost-effectiveness of a biodiversity conservation policy: A bio-econometric analysis of Natura 2000 contracts in forest. *Ecological Economics*, 119.

Holmes, G., et al. (2012), A Comparison of Global Conservation Prioritization Models with Spatial Spending Patterns of Conservation Nongovernmental Organizations. *Conservation Biology*, 26 (4), 602–609.

Independent Evaluation Group (IEG) (2011), *Impact Evaluations in Agriculture: An Assessment of the Evidence*. Washington, DC, World Bank.

Ingram V., et al. (2016), Evaluating the impacts of plantations and associated forestry operations in Africa: methods and indicators, *International Forestry Review* 18(1): 44–55.

Institute for Environmental Studies (IVM) (2005), *Literature review on ex-post assessment of costs to business of environmental policies and legislation*. IVM/Ecologic Institute, Amsterdam/Berlin. http://ec.europa.eu/environment/enveco/ex_post/pdf/literature.pdf.

Jayachandran, S. (2017), Cash for Carbon: A Randomized Controlled Trial of Payments for Ecosystem Services to Reduce Deforestation. *Science*, 357, 6348.

- Jones, K., and Lewis, D. (2015), Estimating the Counterfactual Impact of Conservation Programs on Land Cover Outcomes: The Role of Matching and Panel Regression Techniques, *PLOS One*, 10 (10).
- Joppa, L., and Pfaff, A. (2010), Reassessing the forest impacts of protection, *Annals of the New York Academy of Sciences*, 1185.
- Kronback, L., and Vestergaard N. (2013), Environmental cost-effectiveness analysis in intertemporal natural resource policy: Evaluation of selective fishing gear. *Journal of Environmental Management*, 131, 270-279.
- Kuhfuss, L., and Subervie, L. (2018), Do European Agri-environment Measures Help Reduce Herbicide Use? Evidence From Viticulture in France, *Ecological Economics*, 149.
- Lan, J., and Yin R. (2017), Research trends: Policy impact evaluation: Future contributions from economics, *Forestry Policy and Economics*, 83.
- Lance, P., et al. (2014), How do we know if a program made a difference? A guide to statistical methods for program impact evaluation. Chapel Hill, North Carolina: MEASURE Evaluation.
- Law, E. (2016), Is international conservation aid enough? *Environmental Research Letters*, 11 (2).
- Lawley C and Towe C. (2014), “Capitalized Costs of Habitat Conservation Easements.” *American Journal of Agricultural Economics* 96 (3): 657–72.
- Leppert, G., L. et al. (2018), Impact, Diffusion and Scaling-Up of a Comprehensive Land-Use Planning Approach in the Philippines. From Development Cooperation to National Policies, German Institute for Development Evaluation (DEval), Bonn.
- Le Velly, G., Dutilly, C. (2016), Evaluating Payments for Environmental Services: Methodological Challenges, *PLoS ONE* 11(2), e0149374, <https://doi.org/10.1371/journal.pone.0149374>.
- Luintel H., Bluffstone R.A., Scheller R.M. (2018), The effects of the Nepal community forestry program on biodiversity conservation and carbon storage, *PLoS ONE* 13(6).
- Lund, J.F., et al. (2014), Mixed method approaches to evaluate conservation impact: evidence from decentralized forest management in Tanzania, *Environmental Conservation*, 42 (2).
- Macmillan, D., Harley, D., Morrison R. (1998), Cost-effectiveness analysis of woodland ecosystem restoration, *Ecological Economics*, 27, 313–324.
- Martin, A., et al. (2014), Measuring effectiveness, efficiency and equity in an experimental Payments for Ecosystem Services trial, *Global Environmental Change*, 28.
- McKinnon, M. C., et al. (2015a), Sustainability: Map the Evidence. *Nature Comment*. 528.
- McKinnon, M. C., et al. (2015b), Impact evaluation to communicate and improve conservation non-governmental organization performance: the case of conservation international. *Phil. Trans. R. Soc. London*, B 370.
- McVittie, A., et al. (2014), Something to grouse about? The cost-effectiveness of biodiversity measures in Scotland, Paper presented at the Agricultural Economics Society 88th Annual Conference, 10-11 April 2014.
- Miller, D., et al. (2015), The importance of national political context to the impacts of international conservation aid: evidence from the National Parks of Benin and Niger. *Environ. Res. Lett.* 10.

- Miranda J. et al. (2016), Effects of Protected Areas on Forest Cover Change and Local Communities: Evidence from the Peruvian Amazon, *World Development*, 78, pp. 288–307.
- Miteva DA, Loucks CJ, Pattanayak SK. (2015), “Social and Environmental Impacts of Forest Management Certification in Indonesia”, *PLoS ONE*, 10 (7).
- Miteva, D.A., Murray, B, Pattanayak, S. K. (2015), Do protected areas reduce blue carbon emissions? A quasi-experimental evaluation of mangroves in Indonesia. *Ecological Economics* 119, 127-135.
- Miteva, D., Pattanayak, S., Ferraro, P. (2012), Evaluation of biodiversity policy instruments: what works and what doesn't?. *Oxford Review of Economic Policy*, 28 (1): 69-92.
- Montgomery, C. A., Brown, Jr., G. M. and Adams. D. M. (1994), The Marginal Cost of Species Preservation: The Northern Spotted Owl. *Journal of Environmental Economics and Management*, 26, 111–128.
- Narayan, S., et al. (2016), The Effectiveness, Costs and Coastal Protection Benefits of Natural and Nature-Based Defences. *PLoS ONE* 11(5).
- Nelson, A., and Chomitz, KM (2011), Effectiveness of strict versus multiple use protected areas in reducing tropical forest fires: a global analysis using matching methods, *PLoS One* 6 e22722
- OECD (2018), *Mainstreaming Biodiversity for Sustainable Development*, OECD Publishing, Paris, <https://doi.org/10.1787/9789264303201-en>.
- OECD (2018), *Cost-Benefit Analysis and the Environment: Further Developments and Policy Use*. OECD Publishing, Paris, <https://doi.org/10.1787/9789264085169-en>.
- OECD (2018), *OECD Regulatory Policy Outlook 2018*, OECD Publishing, Paris, <https://doi.org/10.1787/9789264303072-en>.
- OECD (2014), *OECD Framework for Regulatory Policy Evaluation*, OECD Publishing Paris. <http://dx.doi.org/10.1787/9789264214453-en>.
- OECD (2012), “Econometric methods for estimating the additional effects of agri-environment schemes on farmers’ practices, in *Evaluation of Agri-Environmental Policies: Selected Methodological Issues and Case Studies*, OECD Publishing, Paris, <http://doi.org/10.1787/9789264179332-13-en>.
- OECD (2010), *Evaluating Development Co-operation: Summary of key norms and standards*. DERec, DAC Evaluation Resource Centre, 2nd edition. OECD Publishing, Paris, www.oecd.org/derec/guidelines.htm.
- OECD (2006), Outline of principles for impact evaluation, OECD Paris, www.oecd.org/dac/evaluation/dcdndep/37671602.pdf.
- Oinonen, S., et al. (2016), Cost-Effective Marine Protection - A Pragmatic Approach. *PLoS ONE*, 11(1).
- Okumu, B and Muchapondwa, E (2017), Welfare and Environmental Impact of Incentive Based Conservation: Evidence from Kenyan Community Forestry Associations. Economic Research Southern Africa. Working Paper 706. <https://econrsa.org/publications/working-papers/welfare-and-environmental-impact-incentive-based-conservation-evidence>.
- Parker, C., et al. (2012), *The Little Biodiversity Finance Book*. 3rd edition. Global Canopy Programme, UK.
- Pattanayak, S.K., Wunder, S., 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, B. (2012), *Linking monitoring and evaluation to impact evaluation*. No.2. Interaction and Rockefeller Foundation, www.interaction.org/impact-evaluation-notes.

- Peterson, I., Maron, M., et al. (2018), A quantitative framework for evaluating the impact of biodiversity offset policies, *Biological Conservation* 224, 162-169.
- Puri, J. (2017), Using Mixed Methods to Assess Trade-Offs Between Agricultural Decisions and Deforestation. In: Uitto J., Puri J., van den Berg R. (eds) *Evaluating Climate Change Action for Sustainable Development*. Springer, Cham.
- Puri, J., Nath, M, Bhatia, R and Glew, L. (2016), *Examining the evidence base for forest conservation interventions*, Evidence Gap Map Report 4. International Initiative for Impact Evaluation (3ie): New Delhi.
- Puri J., and Dhody B. (2016), Missing the Forests for the Trees? Assessing the Use of Impact Evaluations in Forestry Programmes, In: Uitto J., Shaw R. (eds) *Sustainable Development and Disaster Risk Reduction. Disaster Risk Reduction (Methods, Approaches and Practices)*. Springer, Tokyo.
- Richerzhagen, C., et al. (2016), *Why We Need More and Better Biodiversity Aid*. Briefing Paper 13. German Development Institute, www.die-gdi.de/uploads/media/BP_13.2016.neu.pdf.
- Rodríguez-Rodríguez, D., Rees, S., Rodwell, L. and Attrill, M. (2015), “IMPASEA: A methodological framework to monitor and assess the socioeconomic effects of marine protected areas. An English Channel case study”. *Environmental Science & Policy*, 54, pp.44-51.
- Sabet, S., and Brown, A. (2018), Is impact evaluation still on the rise? The new trends for 2010-2015, *Journal of Development Effectiveness*, 10 (3).
- Samii, C., Lisiecki, M, Kulkarni, P, Paler, L and Chavis, L, 2015. *Payment for environmental services for reducing deforestation and poverty in low- and middle-income countries: a systematic review*, 3ie Systematic Review 17. London: International Initiative for Impact Evaluation (3ie).
- Sills, E., and Caviglia-Harris, J. (2015), Evaluating the long-term impacts of promoting “green” agriculture in the Amazon, *Agricultural Economics*, 46 (1).
- Sills, E., Herrera D, Kirkpatrick AJ, Brandão A, Jr., Dickson R, Hall S, et al. (2015) Estimating the Impacts of Local Policy Innovation: The Synthetic Control Method Applied to Tropical Deforestation, *PLoS ONE* 10 (7).
- Sims, K., and Alix-Garcia, J. (2016), Parks versus PES: Evaluating Direct and Incentive-based Land Conservation in Mexico, www.bioecon-network.org/pages/18th_2016/Sims.pdf.
- Somanathan, E., Prabhakar R, Singh Mehta BS (2009), Decentralization for cost-effective conservation. *Proc Natl Acad Sci USA* 106:4143–4147.
- Sutherland W., et al. (2017), *What Works in Conservation 2017*, Cambridge, UK, Open Book Publishers, www.conservazionevidence.com/pdf/What-Works-in-Conservation-2017.pdf.
- United Nations Evaluation Group (2016), *Norms and Standards for Evaluation*. New York, UNEG.
- Verissimo, D., et al. (2018), Measuring the impact of an entertainment-education intervention to reduce demand for bushmeat, *Animal Conservation*, 21 (4).
- Vianna, A.L.M. and Fearnside, P.M. (2014). Impact of community forest management on biomass carbon stocks in the Uatumã Sustainable Development Reserve, Amazonas, Brazil. *Journal of Sustainable Forestry*, 33(2), pp.127–51.
- Villalobos L., Coria, J., Norden, A. (2018), Has Forest Certification Reduced Forest Degradation in Sweden?, *Land Economics*, 94 (2).

Vincent, J. R. (2016). "Impact Evaluation of Forest Conservation Programs: Benefit-Cost Analysis, Without the Economics", *Environmental and Resource Economics*, 1-14.

Waldron, A., Mooers, A.O., Miller, D.C., Nibbelink, N., Redding, D., Kuhn, T.S., Gittleman, J.L. (2013), Targeting global conservation funding to limit immediate biodiversity declines. *Proc Natl Acad Sci USA*, 110 (29).

White, H. (2009), Theory based impact evaluations: Theory and practice. International Initiative for Impact Evaluation, Working Paper 3, New Delhi, India.

World Bank (2009), Making Smart Policy: Using Impact Evaluation for Policy Making – Case Studies on Evaluations that Influenced Policy, No. 14, http://siteresources.worldbank.org/INTISPM/Resources/383704-1146752240884/Doing_ie_series_14.pdf.

ANNEX 1: Overview of selected impact evaluation methodologies

Experimental methods (randomisation)

Experimental control, particularly through random assignment, makes the estimation of counterfactual outcomes easier. A randomised controlled trial (RCT) is a study design that randomly assigns units (e.g. participants) to a treatment group or a control group, where the treatment consists of the real implementation of a policy or programme, and the control group serves as the counterfactual.

Due to possible ethical considerations regarding which units to exclude from a treatment, Gertler et al. (2016) states that randomised assignment can be used as a programme allocation rule in one of two specific scenarios:

(1) When the eligible population is greater than the number of programme spaces available. When the demand for a programme exceeds the supply, a lottery can be used to select the treatment group within the eligible population.

In the case of environmental policies, such a study design could be used to evaluate a Payments for Ecosystem Services (PES) programme, or another form of a subsidy such as an AES.

(2) When a programme needs to be gradually phased in until it covers the entire eligible population. Gertler states: When a programme is phased in, randomisation of the order in which participants receive the programme gives each eligible unit the same chance of receiving treatment in the first phase or in a later phase of the programme. As long as the last group has not yet been phased into the programme, it serves as a valid comparison group from which the counterfactual for the groups that have already been phased in can be estimated. This setup can also allow for the evaluation to pick up the effects of differential *exposure to treatment*: that is, the effect of receiving a programme for more or less time.

Despite its advantages, the use of experimental design is rare in environmental programmes (Ferraro, 2012). Studies that do exist have more commonly been applied in the energy and water domains (Ferraro and Hanuauer 2014). Box 2.1 summarises an example of an RCT study applied in the context of biodiversity, to evaluate a forest PES programme¹⁹.

¹⁹ Martin et al. (2014) used an RCT to assess PES and policy-regulating mechanisms in Rwanda. Vianna and Fearnside (2014) use an RCT in Brazil to evaluate the effect of decentralised forest management on number of trees damaged per harvested tree and on carbon stocks.

Box A.1. Applying a randomised control trial to evaluate a PES

The study objective was to evaluate a PES programme in Uganda that offered forestry owning households annual payments of 70,000 Ugandan shillings if they conserved their forest. 121 villages with private forest owners were selected for the RCT study. After a baseline survey was undertaken, subcounty-level public lotteries were conducted to choose the 60 villages that were in the treatment group with the rest serving as the control group. The indicator used was tree cover (thus enabling evaluation of both deforestation and forest degradation). The authors found that tree cover declined by 4.2% during the study period in treatment villages, compared to 9.1% in control villages, and that there was no evidence of leakage to nearby land.

Source: Jayachandran et al., (2017), *Science*, 357, pp. 267–273.

Quasi-experimental methods

Matching

With matching, the control group is constructed in order to make it resemble as much as possible the treatment group, on the basis of *observed characteristics*. If resemblance is satisfactory, the outcome observed for the matched group approximates the counterfactual, and the effect of the intervention is estimated as the difference between the average outcomes of the two groups. The method of matching gets rid of selection bias due to observables. It is not able to address bias caused by un-observables.

For example, if the objective is to evaluate the impact of listing and funding species under an environmental protection act in terms of species recovery, a control group of species can be selected that has similar characteristics (e.g., level of endangerment, biological characteristics, political influences, scientific knowledge, and advocacy) (Ferraro and Pattanayak, 2006). Similarly, in the context of an agri-environment scheme, treated farmers can be compared to similar untreated farmers that have the same observed characteristics (e.g., same education level, age, farm structure and equipment but who differ in treatment status). It is then possible to form the difference in outcomes between the treated farmer and his twin. By doing this for all treated farmers and taking the average of these differences, the average causal effect of the AES on participating farmers (i.e. the average effect of treatment on the treated - ATT) can be estimated (OECD, 2012).

Matching methods gives an unbiased estimate of the ATT under three assumptions: 1) that the treatment received by one does not affect outcomes for another; 2) that there are no unobserved characteristics; and 3) that for each participant there exists at least one “twin” nonparticipant having the same observed characteristics (OECD, 2012). Stated otherwise, matching assumes that similarity in the observed characteristics translates into similarity in unobservable characteristics, correlated with the outcome and the biodiversity policy assignment, or that such unobservables are negligible sources of bias (Miteva, 2012).

Matching methods have been applied to the evaluation of a wide range of environmental policies including farmland conservation (Mezzatesta et al., 2013; OECD, 2012), eco-certification (Blackman and Naranjo, 2012), integrated conservation and development projects (Weber et al. 2011), payments for ecosystem services (Arriagada et al. 2012), and ecosystem conservation (Joppa and Pfaff, 2012) (see Ferraro and Hanauer 2014).

Difference-in-differences

There may be systematic differences between the treatment and non-treatment groups even after conditioning on observables. There may also be unobservables, as mentioned earlier, such as motivation or environmental awareness that may influence e.g., a farmer's decision to participate in an AES. Difference-in-difference (DID) designs measure the impact of a policy by the difference in the before–after change in the outcomes. The impact is estimated by computing a double difference: one over time (before–after) and one across subjects (between beneficiaries and non-beneficiaries). DID assumes that any unobserved differences (i.e. systematic biases) are linear and time-invariant and can hence be removed by taking the difference in the outcomes before and after the policy. Simply observing the before and after change in the treatment group is not sufficient as there may be other factors likely to influence the outcome over time. Simply comparing the treatment and control group is also not sufficient. The first difference (for the treatment) controls for factors that are constant over time, since the same group is compared to itself. The second difference is on the comparison group (Gertler et al. 2016). This is illustrated in Figure 2.2 where the DID impact is $= (B - A) - (D - C) = (B - E)$. This method is less complex than other methods, and requires longitudinal aggregate data on policy outcomes for beneficiaries and non-beneficiaries, collected before and after the intervention.

Figure A.1 The difference-in-difference method

Source: Adapted from Gertler et al. (2016)

Box A.2 Using DID to evaluate impact of PES on forest cover in Chiapas Mexico

Using Difference-in-Difference (DID) methods, the study examines the effectiveness of Mexico's national programme of payments for biodiversity conservation, focusing on avoided forest loss as a proxy outcome. The analysis is performed comparing grid cells spread over two municipalities in Chiapas, which involved identifying suitable control groups with good statistical properties in a geographically constrained area that is subject to high deforestation pressure. They examine programme implementation in a geographically limited area where territories have relatively similar historical trends, land use patterns, and potentially similar access to information and to infrastructure. Using data on land cover in 2007 and 2013, they estimate that the additional conservation represents between 12 and 14.7 percent of forest area enrolled in the programme in comparison to control areas.

Source: Costedoat et al. 2015.

Instrumental variables

The instrumental variables (IV) method produces credible estimates of the impact of the policy when the exposure to the policy is to a certain degree determined by an “external force” that does not affect the outcome of the policy directly, but only indirectly, through its influence on exposure. It can help evaluate programmes with imperfect compliance, voluntary enrolment, or universal coverage (Gertler et al. 2016). With IV, the causal effect is estimated by measuring how the outcome varies with the portion of the total variation in the treatment explained by variation in the instrumental variable. For example, if PAs are more likely to be assigned where endemic mammals are present, but the presence of endemic mammals only affects deforestation rates through its effect on the likelihood of a parcel's protection, then the presence of endemic mammals can be used as an ‘instrument’ to identify a causal effect of PAs on deforestation (Miteva et al. 2012). See Amin et al (2015) for an application. In practice, it is often hard to find instruments that are both strong (related to the intervention) and valid (unrelated to the outcome).

ANNEX 2: Inventory of impact evaluation and cost-effectiveness analysis of biodiversity instruments

Table A 2.1

Policy Instrument	Location	Continent	Scope	Unit	Method	Impact	Cost	Reference
Agri-Environment Scheme	France	Europe	French farms 2003-2005	ha of cover crops planted	Difference-in-difference (DID)/matching	Only farming indicators, but strict AES are more cost effective than less strict AES on account of windfall effects	Included	OECD 2012
Agri-Environment Scheme	U.K.	Europe	Cereal farms in the UK Farm Business survey	Farmer behaviours with respect to intensity, productivity and structure of production	Matching	Only economic indicators; no assessment of environmental impact. Found AES to effectively influence individual producer behaviour with respect to intensity, productivity and structure of production	Not included	Sauer et al 2012
Agri-Environment Scheme	France	Europe	Impact on pesticide use based on Viticulture in southern France, 2011 and 2012	Quantity of pesticides used	Matching	Quantity of herbicides used by participants in the scheme in 2011 ranges from 38 to 53% below what they would have used without the scheme, between 42 and 50% in 2012	Not included	Kuhfuss and Subervie 2018
Agri-Environment Scheme	UK, Spain, Italy, France, Germany	Europe	2003-2006	Changes in crop number, fertiliser, and crop protection expenditure per hectare and share of grassland	Matching and DID	The effects of the AESs adoption largely depend on the share of the agri-environmental payment on farm revenue. If this share is larger than 5%, participation in AESs is effective in promoting greener farming practices in all countries but Spain	Included	Arata and Sckokai 2016
Biodiversity measures	Scotland	Europe	Scottish Rural Development Programme and similar schemes	GBP/unit of effectiveness	Cost-Effectiveness Analysis (CEA)	Different actions' cost effectiveness varied between species and habitat from GBP3,286 to 4 million per unit	Included	McVittie et al 2014

Blacklisting	Brazil	Latin America	Legal Amazon	% deforestation rate	Matching	Blacklisting districts led to a significant reduction in deforestation rates, mostly likely due to better field enforcement and individual farmers being forced to register land through the reduced availability of rural credit	Not included	Cisnero et al 2015
Community-based forest management	Ethiopia	Africa	Gera regional forest priority area	% change in forest area	Generalised Least Squares	Community forest area have an initial surge of deforestation in year 1 followed by a significant increase of forest in subsequent years, which a cumulative rate change of 4.8 percentage points when compared to other areas	Not included	Takahashi and Todo 2012
Community-based forest management	Madagascar	Africa	Community forest management between 2000 and 2005	% deforestation rate	Matching	Community forest management makes no difference to deforestation, unless it prevents all forms of commercial activity.	Not included	Rasolofoson et al 2015
Community-based forest management	India	Asia	Uttarakhand, state and village forests	% crown cover in forests	CEA, Matching	State forests are associated with 7-9 times higher costs than village forests, decentralisation could be economically efficient in India	Included	Somanathan et al 2009
Community-based forest management	Kenya	Africa	Plantation Establishment and Livelihood Improvement Scheme (PELIS), Mau forest conservancy	Forest cover	Matching	Impact of PELIS on forest cover estimated at between 5.53% and 7.94%	Not included	Okumu and Muchapondwa 2017

Community-based forest management	Nepal	Asia	National	Seedlings (species and count), saplings (species, count and diameter at breast height), trees (species, count, dbh and height), the environment (slope, altitude, aspect, canopy cover, soil erosion) and NDVI.	Matching	Community Forestry Programme had a significant positive effect on biodiversity	Not included	Luintel et al 2018
Conservation easement	Canada	North America	Southwest Manitoba	Waterfowl habitat suitability and others	Matching	Landowners have been adequately compensated and conservation agencies have successfully secured habitat at risk of conversion	Included	Lawley and Towe 2014
Conservation spending	Global	Multiple	Conservation spending	Index of biodiversity decline, USD	Zero inflated continuous regression modelling	Conservation spending is effective at reducing biodiversity decline, but its impact reduces with increasing development pressures	Included	Waldron et al 2017
Eco-certification	Ethiopia	Africa	Coffee, shade coffee certification	% deforestation rate	Matching	Certification reduces deforestation by 1.7 percentage points	Not included	Takhashi and Todo 2013
Eco-certification	Colombia	Latin America	Coffee Santander province, rainforest coffee	ha	Matching	Certified area had larger forest patches and more connected area of forest	Not included	Rueda et al 2015
Eco-certification	Costa Rica	Latin America	Coffee, Organic certification	Adoption of negative and positive (for the environment) farming practices	Matching	Certification reduces environmentally negative farming practices (eg reduced chemical inputs)	Not included	Blackman and Naranjo 2012
Eco-certification	Kalimantan, Borneo, Indonesia	S.E. Asia		% tree cover, % edge effects, number of active fires	Matching	Certification reduced deforestation by 5 percentage point, 31% reduction in air pollution, 32% reduction in respiratory diseases and minor reduction in malnutrition (1 person)	Not included	Miteva et al 2015

Eco-certification	Sweden	Europe	Non-industrial private forests. Forest plots randomly selected for ground inspection by the Swedish Forest Agency	% rate, ha and number/ha. Multiple measures: Compliance rate, area cleared (ha, and %) and trees and high stumps left/ha	Matching	Certification has not halted forest degradation and has not improved any environmental outcomes	Not included	Villalobos et al 2018
Endangered Species Management	New Zealand	Oceania	Kokako management	NZD/pair/site	CEA	Cost effectiveness varied between sites from NZD 42,976 - 60,094 per pair	Included	Fairburn et al 2004
Enhanced land-use planning	Philippines	S.E. Asia	National, implemented over 10 years		Mixed methods with matching	Limited to moderate impacts in the field of sustainable natural resource management: increase in number of PAs and of conservation projects such as tree planting though no measurable effect on actual change in land use reported by households.		Leppert et al. 2018
Entertainment-education intervention	Tanzania	Africa	Northern Tanzania, radio show to reduce demand for bushmeat	Before and after survey on consumer behaviour	Matching	No differences in outcomes between the treatment and control groups, and thus no evidence of the intervention achieving its initial goals	Not included	Verissimo et al 2018
Farmland conservation	USA	North America	Ohio, Great Miami River watershed	ha of land under certain management practices	Matching	Only farming indicators, found some practices to be more cost effective than others	Not included	Mezzatesta Newburn and Woodward 2013
ICDP	Brazil	Latin America	Tapajós National Forest	% change in forest cover	Matching and DID	ICDP had positive impacts on household income but not assets, livelihood portfolios or forest conservation	Not included	Bauch et al 2014
ICDP	Indonesia	S.E. Asia	ICDP around Kerinci Seblat NP	% deforestation	Matching	ICDP had no impacts on deforestation rate, it was severely undermined by land tenure issues and poor enforcement	Included	Linkie et al 2008

International conservation aid	Sub-Saharan and Southern Africa	Africa	42 African countries	% deforestation rate	OLS regression	Increase in conservation aid is associated with a small but statistically significant increase in deforestation at a 2 year time lag	Included	Bare et al., 2015
Law Enforcement	Brazil	Latin America	Field-based enforcement in legal amazon in Brazil	ha of deforestation	Matching	Inspection effective at deterring large scale infractions, but not small scale	Included	Börner et al 2015
Mapping, monitoring and other activities	Brazil	Latin America	Paragominas state local initiative	km ²	Synthetic control measure	Paragomas successfully reduced its rate of deforestation	Not included	Sills et al 2015
Moratorium	Indonesia	S.E. Asia	National	% deforestation rate, ha, MtCO ₂ e	Matching	Moratorium reduced emissions and would be more effective if more widely implemented. Also estimate a break even carbon price for REDD projects to achieve the NDC required reduction	Not included	Busch et al 2015
PES	Rwanda	Africa	Nyungwe National Park	Number of instances of human activity found inside park boundary	Matching	PES has some impact on activity but there was a lack of true controls, PES did not significantly impact social equity	Not included	Martin et al 2014
PES	Uganda	Africa	Forest (western Uganda)	ha (change in land area covered by trees)	RCT and compared cost of PES with value of delayed CO ₂ emission reductions	Tree cover declined by 2-5% in treatment villages compared to 7-10% in control villages. No evidence of leakage	Included	Jayachandran et al 2017
PES	Ecuador	Latin America	63 farmers enrolled in Socio Bosque programme	% of forest lost per year	Matching	Reduced average annual deforestation by 0.4-0.5%	Not included	Jones et al 2017
PES	Costa Rica	Latin America	Sarapiquí region	ha	Matching and DID	PES scheme led to ~10ha increase in forest cover in PES enrolled farms	Not included	Arriagada et al 2012

PES	Costa Rica	Latin America	National PSA (Pago por Servicios Ambientales) scheme	ha	Matching	0.2% annual deforestation avoided. Low clearance rate anyway and non-random enrolment limited impact	Not included	Robalino and Pfaff 2013
PES	Mexico	Latin America	Federal forest protection programme	% deforestation from normalised difference vegetation index (NDVI)	Matching	PES reduced deforestation rates by about 50%, but when slippage is accounted for the actual reduction is closer to 4%	Not included	Alix-Garcia et al 2012
PES	Mexico	Latin America	Forest in Chiapas	ha	Matching	Additional 12-14.7% forest area enrolled	Not included	Costedoat et al 2015
PES	Mexico	Latin America	South Yucatan	% deforestation in forested polygons	Matching	PES effects ceased after land was removed from programme, leakage effects are shown	Not included	Le Velly et al 2017
PES	Mexico	Latin America	Properties enrolled in National Hydrological services PES programme	% deforestation	Matching	Small reduction in deforestation as a result of PES, but evidence of leakage	Not included	Alix-gracia et al 2012
PES	Mexico	Latin America	PSAH programme	% forest cover change and change in NDVI at randomly selected points	Matching	Reduction in land cover loss by 40-51%, and has a small alleviating effect on poverty. Effects are spatially heterogeneous	Included	Alix-Garcia et al 2015
PES	Cambodia	S.E. Asia	The birds nest programme, northern Plains	% of nest successfully fledging	Marching	3 of 4 species had population increases as a result of the programme	Included	Clements et al 2013
Protected area	Africa	Africa	Protected and unprotected important bird areas (IBAs)	Rate of land conversion	Matching	Land conversion rates in protected IBAs were 58% lower than in unprotected IBA. No evidence of leakage	Not included	Beresford et al 2013
Protected area	Madagascar	Africa	Protected areas across 2 time periods 1990-2000 and 2000-2010.	% deforestation rate	Matching	PA reduced deforestation rates in humid forest (0.6-0.2%/yr), dry forest (0.7-0.5%/yr) and spiny forest (1.1-0.8%/yr)	Not included	Eklund et al - 2016

Protected area	France	Europe	Natura 2000 forests	Biodiversity index value	CEA	Complicated, but suggests costs of interventions do not scale with conservation targets. More ambitious targets would be relatively cheaper	Included	Hily et al., 2015
Protected area	Russia	Europe	12 PA in European Russia	Deforestation in 30m ² plots	Matching	Reduced forest disturbance by ~2 percentage points, slightly more in 2000-2005	Not included	Wendland et al 2015
Protected area	Brazil	Latin America	Legal Amazon	km ²	Spatial autoregression and instrumental variable modelling	PA have reduced deforestation even when spill over effects are accounted for	Not included	Amin et al 2014
Protected area	Brazil	Latin America	Acre State PAs, Brazil	km ²	Matching (covariate)	Protection reduced deforestation in sustainable use area (~1-2%) but not other types of protection	Not included	Pfaff et al 2014
Protected area	Brazil	Latin America	Legal Amazon, Brazil	pixels (0.0081km ²)	Matching (covariate)	Protection reduced deforestation around 2% 2000-2004 and <2% 2004-2008, newer PA avoided slightly less	Not included	Pfaff et al 2015a
Protected area	Brazil	Latin America	Legal Amazon, Brazil, split into arc of deforestation and non arc	pixels (0.0081km ²)	Matching (covariate)	Protection reduced deforestation, but varies between location, time period and type of PA	Not included	Pfaff et al 2015b
Protected area	Brazil	Latin America	PA in the Cerrado	% deforestation	Matching	PAs were broadly effective at reducing deforestation	Not included	Carranza et al 2014
Protected area	Chile	Latin America	Cordilleran Protection Area (CPA) for conservation of highly threatened species; the Huemel (<i>Hippocamelus bisulcus</i>)	Change in population index, returns from land use CL\$	CEA	Current management was not environmentally or economically efficient	Included	Saldarriaga-Isaza (2007)
Protected area	Costa Rica	Latin America	All PA Costa Rica	3ha plots	Matching	11% of protected plots (~54,000-60,000ha) would have been deforested without protection	Not included	Andam et al 2008

Protected area	Costa Rica	Latin America	All PA Costa Rica	3ha plots	Matching	18.3% of land in PA experienced regrowth as a result of protection before 1979. 6.4% for lands protected since 1981	Not included	Andam Ferraro Hanuaer 2013
Protected area	Costa Rica	Latin America	All PA Costa Rica	Pixels	Matching	Protection reduces deforestation rate by ~2%	Not included	Pfaff et al 2009
Protected area	Costa Rica	Latin America	PA in Costa Rica	% deforestation, ha	Matching	Protection resulted in about 8% avoided deforestation (1993-1998) ~19,000ha, and 11% avoided (1997-2005) ~25,000ha	Not included	Andam et al 2007
Protected area	Ecuador	Latin America	Tropical Andean Forests, PA gazetted 1990-2008	ha	Matching	5% deforestation avoided	Not included	Cuenca et al 2016
Protected area	Guatemala	Latin America	Strict protection and mixed use areas in the Maya biosphere reserve	% deforestation	Matching	Both mixed use and strict protection reduce rate of deforestation	Not included	Blackman 2015
Protected area	Mexico	Latin America	PA system	Not specified	Matching	2.9% deforestation avoided	Not included	Pfaff et al 2014a
Protected area	Mexico	Latin America	Mexico PA 1990-2000	% deforestation	Matching	No impact of PA at national scale; at regional scale PA increased or reduced clearance. Newer better-funded PAs were more effective	Not included	Blackman et al 2015
Protected area	Panama	Latin America	PA system (pre 1992 gazettelement)	% rate of deforestation	Matching	PA has some impacts, but drivers of deforestation and PA impacts shift over time	Not included	Haruna et al 2014
Protected area	Peru	Latin America	PA, private conservation concession, and indigenous territories in the Peruvian amazon	% rate, ha	Matching	All three governance regimes reduced deforestation and degradation when compared to unprotected land, but not logging or mining concessions. Also CCs and IT were more effective than PA	Not included	Schleicher et al 2017

Protected area	Peru	Latin America	PA pre-2000	forest cover change (both deforestation and degradation), 30m ² plot of land	Matching	8% forest disturbance avoided, no impact on poverty	Not included	Miranda et al 2016
Protected area	Developing Countries	Multiple	PA up to 2008	Fire in forest pixels	Matching	PA reduce frequency of fires in Latin American and more recent PA achieve the same thing in Africa	Not included	Nelson and Chomitz 2009
Protected area	Global	Multiple	Global PA	Change in 1km ² pixels from natural to modified land cover	Matching	PA prevent land cover change, but impact is halved when controlling for confounding variables	Not included	Joppa and Pfaff 2011
Protected area	Bolivia, Costa Rica, Indonesia, Thailand	Multiple	PA systems in country, Indonesia limited to Sumatra	Forest loss per pixel (100m ² Bolivia, 3ha Costa Rica, 1km ² Indo, 30m ² Thailand)	Matching	PA reduced deforestation but impacts of level of protection are globally heterogeneous	Not included	Ferraro et al 2013
Protected area	Costa Rica and Thailand	Multiple	PA Thailand and Costa Rica	Deforestation in 3ha plots	Matching and LOESS regression	Complicated, different conditions are associated with poverty reduction and reduced deforestation in PAs	Not included	Ferraro et al 2011
Protected area	Indonesia	S.E. Asia	PA	% deforestation in 6.9km x 6.9km grid cells	Bayesian spatial modelling (autologistic and von-Thunen spatial autoregressive)	PA do not prevent deforestation in Indonesia	Not included	Brun et al 2015
Protected area	Indonesia	S.E. Asia	Sumatra	% deforestation in 10km grid cells	Matching	PA prevented deforestation compared to non protected region (reduced ~17 percentage points)	Not included	Gaveau et al 2009
Protected area	Indonesia	S.E. Asia	Sumatra	km ²	Matching	PA better than conversion areas but not production areas at preventing deforestation	Not included	Gaveau et al 2012
Protected area	Indonesia	S.E. Asia	MPA and species management PA spanning mangroves and blue carbon emissions	ha of mangroves, mtCO _{2e}	Matching and DID	MPAs reduced mangrove loss by about 14,000ha and avoided ~13mtCO ₂	Not included	Miteva et al 2015

Protected area	Thailand	S.E. Asia	Doi Inthanon, Ob Luang, Thailand	All within 5kmx5km grid, forest cover(%), forest patch size(km ²), and density of cleared patches(#patches/km ²)	Matching	19% more forest in Protected Areas	Not included	Sims et al 2014
Protected area and logging ban	China	Asia	Logging ban, PA, and sacred forest in Yunnan	km ²	Matching and panel regression	PA protected old growth forest but not forest as a whole. Logging ban protected all forest but led to an increase in logging in sacred forest areas	Not included	Brandt et al 2015
Protected area and PES	PES in Ecuador; PA in Russia	Multiple	NA	% deforestation rate	FE panel (DID)	PES in Ecuador resulted in a 42-72% reduction in deforestation, in Russia there was 0-36% reduction. Method of calculation makes a big difference	Not included	Jones and Lewis 2015
Protected area and PES	Cambodia	S.E. Asia	Two PAs, Kulen Promtep & Preah Vihear	% deforestation rate	Matching and DID	PA reduced deforestation by 60% and PES by another 50% on top of PA effect	Not included	Clements et al 2015
Protected area and PES	Costa Rica	Latin America	National	% deforestation rate	Matching	Greater additionality can be achieved when the 2 interventions are separate	N/A	Robalino, 2015. Evaluating Interactions of Forest Conservation Policies
Protected area and PES	Mexico	Latin America	Monarch butterfly habitat	ha of forest cover (deforestation and disturbance)	Matching	Protection has avoided 200-710ha deforestation in high quality habitat, but less in other forest	Not included	Honey-Roses et al 2011
Protected area and PES	Mexico	Latin America	National	% deforestation rate	Matching, CEA (or combined IE w/C)	Both PES and PA avoided similar amounts of deforestation (20-25%), PES was associated with small improvement in livelihood and PAs were livelihood neutral	Not included	Sims and Alix-Garcia, 2016

REDD+ projects	Nepal	Asia	REDD+ projects in Dikha, Gorkha and Chitwan	tC/ha, incidence of forest fires, timber extraction and encroachment	Matching and DID	Reduced forest fires, and timber extraction, decline in use of firewood in REDD communities. No change in forest carbon though	Not included	Sharma et al 2015
REDD+ projects	Brazil	Latin America	PES REDD pilot project Brazil	ha	DID	Reduced forest loss on enrolled farms by 50%, extra forest comes at the expense of pasture. Project was cost effective	Included	Simonet et al 2015
Social Marketing campaigns	USA	North America	Cobb County, Atlanta, Georgia for residential water demand/conservation	% change in water use	Randomised experimental design	Social comparison messages were more effective than simple technical information or prosocietal messages. Their effect wanes over time	Not included	Ferraro and Price 2011
Social Marketing campaigns	Philippines	S.E. Asia	Fishery sustainability in Philippines MPA	Ordinal unit of fish biomass (low, medium high)	Matching and DID	Campaign increased compliance, but no effect seen on biomass	Included	Verissimo et al 2018
Technical and other assistance for green agriculture	Brazil	Latin America	Rondonia, Brazil. Membership in Alternative Producers Association to promote agroforestry	Land cover using remote sensing data	Matching and DID	Membership resulted in more diversified production systems, including more land allocated to agroforestry. Members also deforested less of their farms, but difference is not statistically different	Not included	Sills and Caviglia-Harris 2015
UK Biodiversity Action Plan	U.K.	Europe	380 single Species Action Plans (SAPs)	GBP to achieve SAPs	CEA	Spending distribution is highly biased toward vertebrates, but invertebrate projects tend to be more efficient	Included	Laycock et al 2009
Zoning policy	Cameroon	Africa	FMU in Bertoua	% change in forest area	Matching	Deforestation rates were low inside areas zoned for forest protection, communities also perceived an improvement in forest protection	Not included	Bruggeman et al 2015