ANNUAL REVIEWS

Annual Review of Resource Economics

The Effectiveness of Forest Conservation Policies and Programs

Jan Börner,[1] Dario Schulz,[1] Sven Wunder,[2,3] and Alexander Pfaff[4]

[1]Institute for Food and Resource Economics and Center for Development Research, University of Bonn, 53115 Bonn, Germany; email: jborner@uni-bonn.de, dario.schulz@uni-bonn.de

[2]European Forest Institute, 08025 Barcelona, Spain; email: sven.wunder@efi.int

[3]Center for International Forestry Research, Lima 12, Peru

[4]Sanford School of Public Policy, Duke University, Durham, North Carolina 27708, USA; email: alex.pfaff@duke.edu

Annu. Rev. Resour. Econ. 2020. 12:19.1–19.20

The *Annual Review of Resource Economics* is online at resource.annualreviews.org

https://doi.org/10.1146/annurev-resource-110119-025703

JEL codes: N5, Q15, Q23, Q28, Q58, P48

Keywords

deforestation, forest policy, impact evaluation, meta-analysis, systematic review, evidence gap

Abstract

The world's forests provide valuable contributions to people but continue to be threatened by agricultural expansion and other land uses. Counterfactual-based methods are increasingly used to evaluate forest conservation initiatives. This review synthesizes recent studies quantifying the impacts of such policies and programs. Extending past reviews focused on instrument choice, design, and implementation, our theory of change explicitly acknowledges context. Screening over 60,000 abstracts yielded 136 comparable normalized effect sizes (Cohen's *d*). Comparing across instrument categories, evaluation methods, and contexts suggests not only a lack of "silver bullets" in the conservation toolbox, but that effectiveness is also low on average. Yet context is critical. Many interventions in our sample were implemented in "bulletproof" contexts of low pressure on natural resources. This greatly limits their potential impacts and suggests the need to invest further not only in understanding but also in better aligning conservation with local and global development goals.

1. INTRODUCTION

Land-based resources, and natural forests in particular, are under increasing pressure globally due to economic and population growth as well as associated shifts in consumption (Nkonya et al. 2016, Scholes et al. 2018). At the same time, many countries are taking measures to achieve the United Nations' Sustainable Development Goals (SDG), many of which are directly or indirectly linked to land use and management (Vlek et al. 2017). It thus becomes ever more important to design and implement effective policies that coherently interact to provide incentives for sustainable land use and land management (Seymour & Harris 2019).

In moving toward this goal, policy makers have not always taken into account that environmental and conservation policies are only one dimension of the policy mix faced by land users. Often, economic and agricultural policies or infrastructure investments and international commodity trade are more powerful drivers of land-use and land cover change than the regulations and incentives provided by environmental administrations (De Sy et al. 2019, Lambin et al. 2001, Rudel 2017). Such economic and development policies, e.g., roads, land rights, concessions, and credit support, not only have their own direct effects on forests but they also moderate impacts from conservation action. The effectiveness of environmental policies thus often hinges on the alignment of instruments across policy sectors with conflicting goals (Harahap et al. 2017).

Especially in developing and emerging economies, the capacities and institutional frameworks for conservation policy design and implementation exhibit governance gaps that can result in noncompliance with legal regulations or ineffectiveness despite compliance (Abman 2018, de Freitas et al. 2017, Sans et al. 2018). Acknowledging this has led to interest in making use of other capacities, including hybrid approaches to conservation, such as supply chain governance initiatives that involve cooperation between public sectors, private actors, civil society organizations, and local communities (Lambin et al. 2018).

Given the significant amount of public resources and international transfers for conservation implied above, understanding what works is important. Impact evaluation of conservation policy instruments has become a dynamic field of research with rapid methodological development (Baylis et al. 2016). Case-based evaluations and early synthesis studies focus on single policy instruments and suggest that, beyond instrument choice, the effectiveness of conservation policies varies substantially by intervention context, design, and implementation style (Börner et al. 2016). Here, we adopt a multiple instrument perspective in taking stock of the recent literature evaluating nature conservation policies. Our review focuses on forest cover as a primary outcome measure but also briefly touches on cobenefits.

Below, we take a brief look at influences on the development of a still nascent culture of evaluation and learning in conservation. We also compare the relatively recent history of conservation evaluation with the more established field of development evaluation in terms of factors that influence the adoption of evidence-based approaches to policy design and implementation.

In Section 2, we condense existing economic theories on the working principles of environmental policies into a generic theory of change for conservation policies that use incentives, disincentives, and enabling measures (Börner & Vosti 2013). Our proposed theory of change focuses on theoretical impact channels, economic and institutional requirements for effectiveness, goal tradeoffs including through leakage and spillover effects, and behavioral issues. Third, in light of those possibilities, we synthesize the recent literature evaluating forest conservation policies in Section 3 based on a systematic review of counterfactual-based evaluation studies and selected existing review studies. The focus on average impacts is then extended by a discussion of the importance of impact heterogeneity across contexts, for any given policy, including due to interactions with development policies (Section 4).

1.1. The Late Arrival of Impact Evaluation in the Conservation Sector

The environment became an independent policy domain only in the middle of the twentieth century. In the 1950s, for example, Britain and the United States introduced pollution control policies commonly regarded as the beginning of modern environmental policy (Andrews 2006). In contrast to economic development, which gained a solid foothold in the international policy agenda in the aftermath of World War II, many natural resource and biodiversity conservation goals arose in the 1970s and 1980s. Documentation and accountability efforts might then also naturally arrive more recently (Ferraro 2009).

Furthermore, many issues currently on international conservation policy agendas were originally promoted by grassroots and civil society organizations, which still play important roles in policy design and implementation today. Cashore et al. (2006) argued that sustainable product certification and eco-labeling practices emerged as a civil society innovation in response to the slow adoption of conservation policy by national governments. Yet, small nongovernmental organizations (NGOs) may find it harder to operate independent evaluations than do large international organizations (Sen 1987). Correspondingly, only large and international environmental NGOs have recently begun to include rigorous evaluation designs in their conservation programs (Parks 2008).

Conservation tends to target places, whereas most development programs target people. Even when conservation initiatives also target people, such as for payments for environmental services (PES), outcomes have to be verified by measuring land-based flows in ecosystem services or corresponding land-use proxies. This has implications not only for data collection (e.g., specialized remote sensing and data processing capacity requirements) and the spatial targeting of interventions (Wahlén 2014). Fundraising strategies, for example, may rely on images of untouched natural habitats with large mammals and scenic landscapes. Land users in such landscapes often happen to have lower profits, implying lower opportunity costs of conservation, which allows for the enrollment in conservation of larger tracts of land under lower conflict risk than in regions with higher population densities and land-use interests. Conservation interventions may then come to be biased against partially degraded landscapes with dynamic land-use change and thus higher potential to achieve additional conservation. This "high and far" type of location bias has been repeatedly confirmed in evaluations of protected areas (PAs) and some PES initiatives (Börner et al. 2013, Joppa & Pfaff 2009). If so, then conservation programs in nonthreatened places with high environmental value could have painted a too-rosy picture about the additionality, i.e., the incremental environmental impact vis-à-vis the reference scenario without a program. This in turn may have lowered the pressure from donors and civil society and internally in implementing organizations to rigorously evaluate impacts of their action.

Development and environmental policies have thus experienced different levels of priority setting in national and international policy arenas, with development policy clearly dominating in the early twentieth century. The two policy domains also differ in terms of (*a*) the actor constellations that drive their agendas and implementation styles, (*b*) the underlying fundraising mechanisms, and (*c*) the data types and related research competencies required to rigorously evaluate their outcomes. Moreover, and especially in developing countries, environmental policies must often be aligned with development goals, such as poverty alleviation, which may involve effectiveness trade-offs. These differences have likely contributed to the late development of an evaluation culture in the conservation policy domain and may, as we show below, also help to explain the still small number of evaluation studies that report on highly effective interventions in the literature.

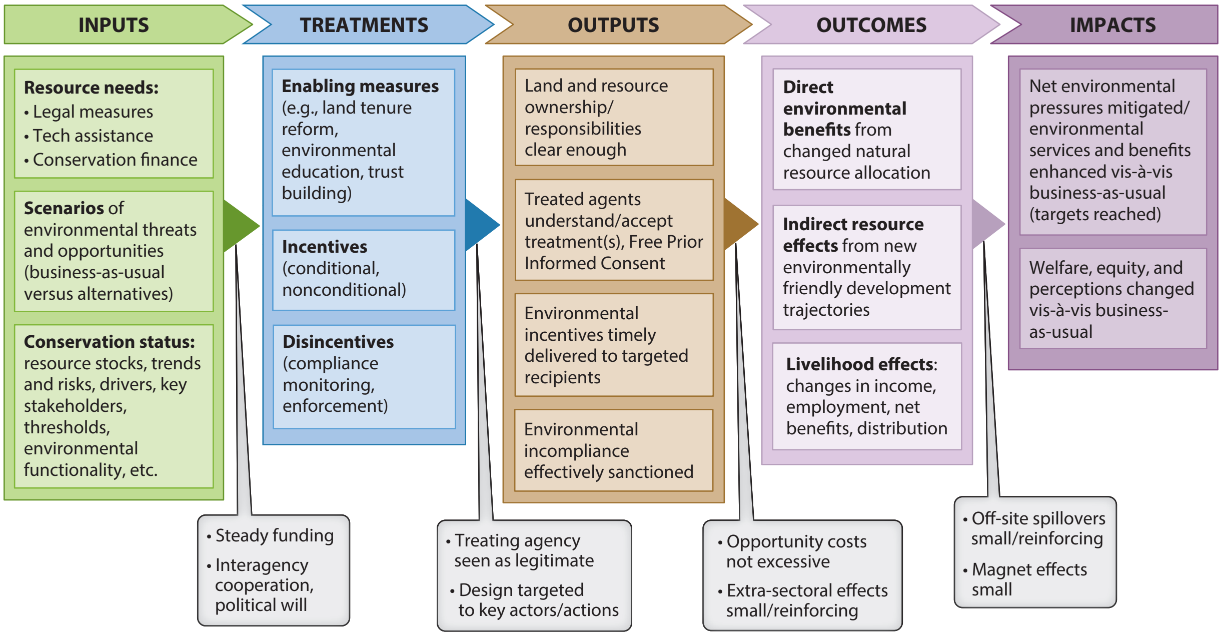

Figure 1

A theory of change for forest conservation initiatives.

2. A THEORY OF CHANGE FOR CONSERVATION

To organize and ease interpretation of the literature reviewed below, this section uses a schematic theory of change (ToC) (Weiss 1997) to lay out multiple theoretical assumptions that typically underlie nature conservation actions. Given the diversity of types and scales of conservation interventions, ranging from tiny integrated community conservation projects to large-scale national PA networks, we necessarily proceed in a relatively generic fashion. **Figure 1** depicts typical ToC reasoning, linking inputs and assumptions to outcomes and impacts at varied scales in space and time.

Reading **Figure 1** from left to right, initial inputs typically include baseline assessments of the state and trends of nature, plus current or potential risks. These often describe the delivery of ecosystem services, the functionality of the ecological system, and thresholds of degradation, beyond which the system can, over a defined time horizon, become altered in ways that affect benefits to people. Scenarios of change implicitly or explicitly identify points of systemic leverage for mitigating threats or generating environmental improvements relative to the perceived business-as-usual scenario. Beyond such knowledge, conservation also requires financial, legal, or technical resources. Finance for conservation often constitutes a key bottleneck.

Input needs depend on the envisaged conservation treatment. **Figure 2** lists types of instruments typically available in the conservation toolbox categorized in terms of incentives, disincentives, and enabling measures (Börner & Vosti 2013). Each aims to influence the private agents' rationales for land- and resource-use decisions toward more environmentally benign outcomes (Vatn 2005).

Disincentives (or "sticks," such as taxes, PAs, logging bans, and law enforcement) restrict or discourage private actions, typically reducing the welfare of the affected agents—to what degree depends on monitoring, sanctions, and compliance (Robinson et al. 2010).

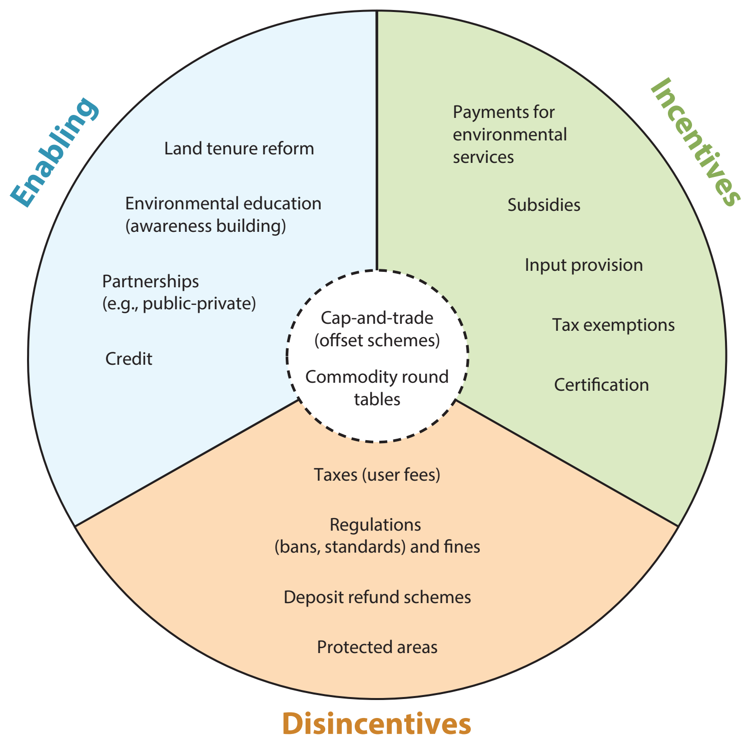

Figure 2

A categorization of conservation interventions. Adapted with permission from Börner & Vosti (2013).

Conversely, incentives ("carrots") typically raise targeted agents' welfare, which could be through a quid pro quo reward conditional upon either an environmentally desirable result (e.g., forest carbon sequestered) or a corresponding land-use proxy (e.g., number of trees reforested). PES (Engel 2016, Wunder 2015) and environmental certification (Blackman et al. 2018) are both examples of voluntarily contracted, yet conditional rewards linked directly to environmental outcomes. Alternatively, subsidies can be used to support changes in livelihood and production toward environmentally more benign land uses, i.e., an indirect or livelihood-oriented conservation approach, such as promoting nondegrading alternative income-generating activities within ICDPs (integrated conservation and development projects) or raising agent incomes to lower dependence on environmental degradation. This approach hopes for win-win outcomes in which key environmental problems are solved while improving livelihoods or welfare of the targeted agents (Ambec et al. 2013).

Lastly, enabling measures may be less linked to agents' outcomes. Some outcomes, such as environmental education, clearly derive from environmental problems (Tilbury 1995). Others, such as land-tenure reform (transfer of property rights and/or consolidation of tenure security), instead change the rules of the game, altering the flows of benefits from natural resource management and/or the degree to which stakeholders can count on those flows. Given longer investment horizons, stakeholders may then preferably opt for longer-term economically profitable land uses (Sunderlin et al. 2008). These land uses may coincide with environmentally desirable scenarios [e.g., tree planting on degraded lands or conservation of native forests (Blackman et al. 2017)] or not [e.g., deforesting land for agricultural conversion (Liscow 2013)]. If enabling measures do not per se link to conservation, it may often be safer to use them jointly with more direct (dis)incentives in integrated policy mixes, so as to guide their impacts in desirable directions.

For any single or multicomponent treatments (second column in **Figure 1**) to translate into desired outputs (third column), it is important that they target key actors and corresponding actions vis-à-vis the points of environmental leverage. Such targeting hinges on the quality of the initial diagnosis of the conservation context. Likewise, the implementing agency needs to be seen as legitimate by the relevant actors, especially when incentivizing voluntary actions, such as adoption decisions by private actors. Free prior and informed consent typically constitutes a necessary condition local acceptance of conservation interventions (Mahanty & McDermott 2013).

Whether the outputs, as short-run treatment results, lead to the desired intermediate outcomes (fourth column in **Figure 1**) in the intervention's target area and time horizon will depend on additional assumptions to hold. First, treatments must have been sufficient to trigger change: If conservation opportunity costs are high, targeted agents may prefer to continue business-as-usual land-use plans. Second, other exogenous factors, such as the macroeconomy or shocks in leading agricultural sectors, can dominate conservation policies when determining these outcomes (Kaimowitz & Angelsen 1998). Hence, whether there is political will to coordinate conservation goals with other policies is a key factor of success, as for instance in Brazil's decade-long (2005–15) dramatic reduction of deforestation (Arima et al. 2014). That said, sometimes total gains are greater than direct environmental benefits from interventions, because indirect effects from other shifts (e.g., reduced deforestation from increased urban labor absorption) are part of gains. The optimal balance between those two is a subject of debate in conservation (Ferraro & Kiss 2002).

Moving from outcomes to longer-term or larger-scale impacts (last column in **Figure 1**)—the treatment's additionality, defined as change in the outcome compared to the business-as-usual scenario—requires further assumptions to hold. For instance, large success in raising local agents' income may lead to "magnet effects" of attracting migrants, whose incremental economic activities can jeopardize the permanence of initial outcome gains (Wittemyer et al. 2008). Likewise, for larger-scale environmental problems exceeding the spatial boundaries of an intervention or a policy's target area (e.g., climate change or global biodiversity extinction), there is a danger of spillovers, including leakage or displacement of threats from the project area to elsewhere. That diminishes the additionality of treatment impacts. Such counteracting leakage effects may well apply with greater force to certain types of scenarios: small projects, elastic commodity supply, flexible supplies of capital and labor, etc. (Pfaff & Robalino 2017).

These conceptual considerations suggest that, in principle, much is known about designing effective conservation policies and programs as well as the potential constraints on impacts. We return to some of the aspects raised above after evaluating the available empirical evidence on the impact of such interventions.

3. SYNTHESIS OF THE RECENT LITERATURE EVALUATING FOREST CONSERVATION INITIATIVES

3.1. Study Selection

To synthesize empirical evidence on the impact of forest conservation initiatives, we created a database of forest cover impact estimates combining different methods. We used a Boolean search string (see **Supplemental Materials**) on three scientific databases to obtain studies with quantitative impact estimates of forest protection measures. We screened 62,637 abstracts resulting from this search using machine learning techniques (Manning et al. 2008, Pedregosa et al. 2011). To gather the initial training sample, we collected quantitative estimates published in existing

Table 1 Overview of reviews and meta-studies on forest conservation policies and programs

Reference[a]	Intervention focus	Methodological notes
Caplow et al. (2011)	REDD+	Systematic review
Porter-Bolland et al. (2012)*	Community-based forest management	Systematic review
Bowler et al. (2012)	Community-based forest management	Systematic review
Geldmann et al. (2013)*	Protected areas	Systematic review
Samii et al. (2014a)*	Decentralized forest management	Systematic review
Samii et al. (2014b)*	Payments for environmental services	Systematic review
Robinson et al. (2010)	Tenure	Systematic review
Macura et al. (2015)*	Governance type	Systematic review
Coad et al. (2015)	Protected areas	Database of management effectiveness
Oldekop et al. (2016)*	Protected areas	Systematic global meta-study
Börner et al. (2016)*	Cross-cutting	Collection overview
Leisher et al. (2016)	Gender	Systematic review
Ojanen et al. (2017)	Property regimes	Systematic review
Lambin et al. (2018)	Supply chain initiatives	Review of recent studies
Snilsveit et al. (2019)	Payments for environmental services	Systematic review
Burivalova et al. (2019)	Cross-cutting (several outcomes)	Systematic review

[a]Studies marked with * were identified prior to data collection for this study and used to create the training sample. Notably, studies published after 2017 were not present in these reviews, whereas studies before 2013 are very likely to be included in at least one of the reviews.

literature reviews and checked them for our relevance and robustness criteria: (*a*) outcome variable must relate to forest cover, (*b*) a counterfactual-based analysis method was used with explicit use of some bias reduction technique, and (*c*) the study was published in a peer-reviewed journal. The literature reviews used to identify training samples are listed in **Table 1**. The studies with observations that met our inclusion criteria made up the training sample ($n = 60$). We iterated six times abstract text mining and classification followed by human expert screening of the 100 studies with the highest relevance score, each time increasing the number of training samples. By personally screening 700 abstracts classified as relevant, we retrieved 167 full texts published between 2001 and 2020 that, based on the abstract, met the inclusion criteria. Finally, we added 17 relevant studies that were referenced in a full text but not yet present in the database.

From the full texts, we extracted quantitative impact estimates as well as information on location, study design, and evaluation method. Where possible, we calculated a normalized effect size (Cohen's *d*) by dividing the treatment effect by the standard deviation of the outcome variable in the control group. As such, Cohen's *d* can be interpreted as the difference between the treated and control groups in terms of standard deviations of the outcome variable. For example, a Cohen's *d* equal to 0.2 implies that approximately 58% of the observations that were subject to a given intervention will exhibit an outcome above the mean outcome of the control group. If standard deviations were not provided in the paper or its supplementary material, we contacted the authors. If we could not obtain the standard deviation of the control group, we used the pooled standard deviation from all observations. We standardized effect sizes to be negative if the underlying effect was a significant reduction in forest cover loss. To avoid bias arising from the diversity of model specifications within one study (i.e., for temporal or spatial subcategories), we took the mean of estimates for the same intervention by the same study in the same country. We conducted an analysis of variance (ANOVA) for intervention outcome, analysis method, and country using the R software environment (by R Core Team; **https://www.R-project.org**). **Supplemental Figure 1** provides a schematic overview of the steps to create the observation database.

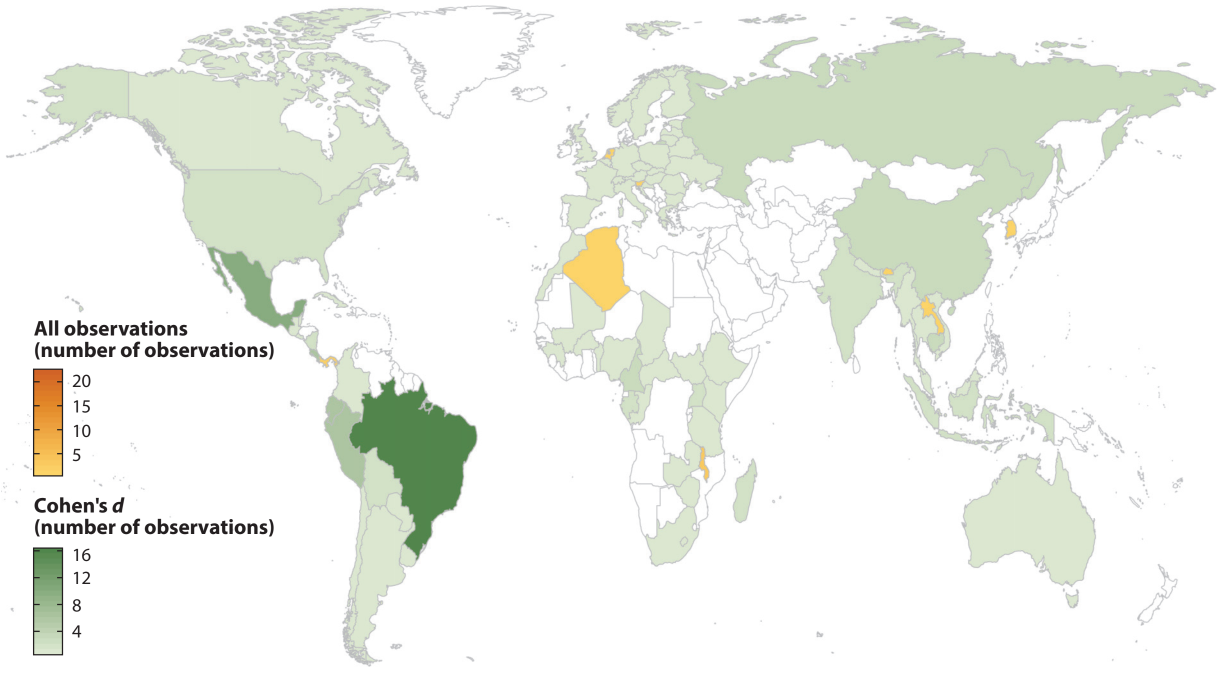

Figure 3

Frequency counts of comparable (*green*) and incomparable (*orange*) effect size observations per country. Cohen's *d* was calculated only for comparable observations. Global studies are not included in this map.

3.2. Characteristics of the Final Sample

After applying the selection criteria, we identified 99 studies reporting counterfactual-based treatment effects of conservation mechanisms. Within these studies, we were able to calculate 198 effect sizes from 51 papers. After taking means as described above, the final number of comparable estimates was 136. One study (Abman 2018) yielded a total of 68 country-specific effect sizes for PAs, but our results are robust to excluding this subsample of PA effect estimates. Most studies focused on specific regions within one country ($n = 37$); national ($n = 10$), cross-country ($n = 3$), and global studies ($n = 1$) are less frequent. **Figure 3** shows the number of weighted observations in each country.

Countries marked in orange color scales in **Figure 3** show the number of forest cover–related estimates for which we were unable to calculate a comparable effect size. Several interventions in Panama and Malawi had been evaluated rigorously, but results could not be included in our meta-analysis due to missing information for the calculation of effect sizes. For Africa, a total of 11 eligible observations could not be included in the analysis because of lacking data, for Latin America 29, for Asia 22, two for Europe, and one for Australia. With only 15 robust studies in total, Africa is the most underrepresented region in our sample, despite the large number of conservation efforts on the continent.

For most countries in North America and Europe as well as for Australia we found two or fewer comparable observations. This is not due to a general lack of evaluation studies, but rather because agri-environmental program evaluations often focus on other indicators than forest cover outcomes. Within the tropics and subtropics, most comparable observations are in Latin America

Table 2 Cohen's *d* observations by intervention type and evaluation method[a]

Intervention type	Evaluation method							
	M	M + DD	FE	Reg.	IV	SCA	RCT	Total
PA	**79** (104)	**2** (3)	**0** (2)	**1** (3)	NA	NA	NA	**82** (112)
PES	**14** (17)	**6** (7)	**0** (2)	**0** (1)	NA	NA	**2** (3)	**22** (30)
DFM	**1** (7)	**0** (1)	**0** (2)	**0** (2)	**2** (3)	NA	NA	**3** (15)
Certification	**1** (4)	**1** (2)	**1** (1)	NA	NA	**3** (3)	NA	**6** (10)
Sustainable use	**7** (8)	NA	**0** (1)	NA	NA	NA	NA	**7** (9)
LTR	**1** (4)	**1** (1)	**1** (1)	NA	**1** (1)	NA	NA	**4** (7)
Governance	**2** (3)	**0** (1)	NA	**1** (1)	NA	**1** (1)	NA	**4** (6)
IPL	**3** (6)	NA	NA	NA	NA	NA	NA	**3** (6)
ICDP	**2** (2)	**0** (1)	**1** (1)	NA	NA	NA	NA	**3** (4)
Voluntary conservation	**2** (2)	NA	NA	NA	NA	NA	NA	**2** (2)
Total	**111** (156)	**10** (15)	**3** (10)	**2** (7)	**3** (4)	**4** (4)	**2** (3)	**136** (201)

[a]Distribution of calculated Cohen's *d* (bold) included in this study and all robust observations (in parentheses) according to intervention (vertical axis) and analysis method used (horizontal axis).

Abbreviations: DFM, decentralized forest management; DD, double-difference; FE, fixed effects regression; ICDP, integrated conservation and development project; IPL, indigenous protected land; IV, instrumental variable; LTR, land titling and reform; M, matching; NA, not applicable; PA, protected area; PES, payments for environmental services; RCT, randomized controlled trial; Reg., other regression; SCA, synthetic control analysis.

($n = 57$), followed by Africa ($n = 25$) and Asia ($n = 18$). Brazil is the country with the highest number of observations ($n = 16$), likely owing to its size, the diversity of conservation mechanisms, and international interest to conserve forests in the Amazon region.

Most studies used remotely sensed land cover pixels as their unit of observation, varying in resolution between 30 and 1,000 meters. Depending on the study, pixels were aggregated to larger units, such as grids, and property or administrative boundaries. Some studies relied on household and forest transect surveys to derive additional biophysical or socioeconomic outcome and control variables. Most studies used some quasi-experimental design to estimate the average treatment effect on the treated (ATT), especially covariate matching (81%; see **Table 2**). In contrast, only two comparable studies applied a randomized controlled trial (RCT) design: Jayachandran et al. (2017) for a PES scheme in Uganda and Wiik et al. (2019) for PES in Bolivia. We have not found studies using regression discontinuity design (RDD) to evaluate forest cover outcomes of conservation policies, although RDD has been used to evaluate socioeconomic outcomes (Alix-Garcia et al. 2018). The synthetic control method was applied to evaluate a combination of public disclosure and rewards at an administrative level (Sills et al. 2015) and certification schemes (Rana & Sills 2018).

PA is the intervention with the highest number of published treatment effects, followed by PES (**Table 2**), although this order reverses when excluding the 68 effect sizes from Abman (2018); PES impacts are covered in detail in another contribution to this volume (Wunder et al. 2020). PAs comprise the largest area of conservation initiatives, and decades of research looked into their effectiveness. Some global studies aggregate over large number of PAs, e.g., Nelson & Chomitz (2011), while many older studies applied a naïve inside-outside comparison method, thus not meeting our selection criteria. In addition, several PA studies did not provide the necessary descriptive statistics to calculate normalized effect sizes. In a few cases, additional information could be obtained by contacting the authors. Only Abman (2018) quantified heterogeneous treatment effects of macrolevel governance characteristics on the effectiveness of conservation initiatives. More recently, PES has received growing attention and substantial funding in the context of REDD+, with different sectors demanding evaluations of its effectiveness.

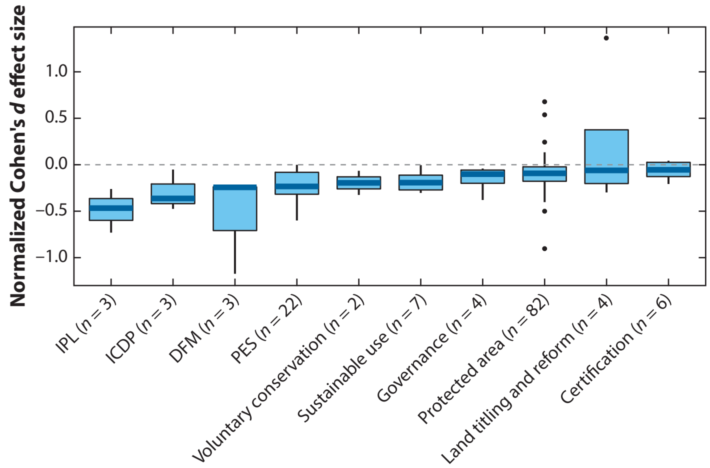

Figure 4

Standardized forest conservation effect sizes (Cohen's *d*) by intervention type ($n = 136$). Boxes show the interquartile range, horizontal bars show the median, whiskers extend to the smallest/largest value no further than 1.5 times the interquartile range from the box, and the dots are outliers. Negative values indicate avoided deforestation, and more negative values imply higher forest conservation effects. Abbreviations: DFM, decentralized forest management; ICDP, integrated conservation and development project; IPL, indigenous protected land; PES, payments for environmental services.

3.3. Effect Size Comparison

We compared the effectiveness of forest conservation mechanisms by means of a normalized effect size (Cohen's *d*; **Figure 4**). According to Sawilowsky (2009), Cohen's *d* ranging between 0.01 and 0.2 can be considered small effects. However, similar ranges were reported for interventions in other sectors, such as for education interventions in developing countries (Petrosino et al. 2012). The highest effect size comes from Heltberg (2001) for a decentralized forest management scheme (governance) in India, using an instrumental variable approach. Another very high estimate originates from an instrumental variable approach (Liscow 2013), suggesting that land titling (not specifically designed for the purpose of conservation) increased deforestation by 13%. The outliers for PA in **Figure 4** and the matching subgroup in **Figure 5** come from Abman (2018), Arriagada et al. (2011), Nelson & Chomitz (2011), and Paiva et al. (2015). For the robust comparison of group means, we removed the two instrumental variable outliers and repeated the entire analysis without the observations from Abman's study; **Figures 4–6** and statistical results remained essentially unaffected. The reader should keep in mind, however, that some subgroup sample sizes are very low and do not allow strong conclusions about the separate roles of context, instrument choice, or evaluation methods in determining effect sizes.

Designation of indigenous areas is on average the most effective intervention in our sample. Using analysis of variance, we found that the difference to both PES and PAs was significant at the 95% level of confidence (**Figure 4**). The estimates for demarcation of indigenous territories require cautious interpretation because of the unique legal status of these lands in Latin America. All observations are from Brazil, where indigenous land differs from nature reserves with sustainable use. On other continents, indigenous communities often have privileged access to areas legally defined as nature reserves, merging the effectiveness of indigenous demarcation into the categories "Sustainable Use" or "PAs," the latter being predominantly strict protection. In contrast, land titling in general has ambiguous effects on forest cover (see Section 2). PAs

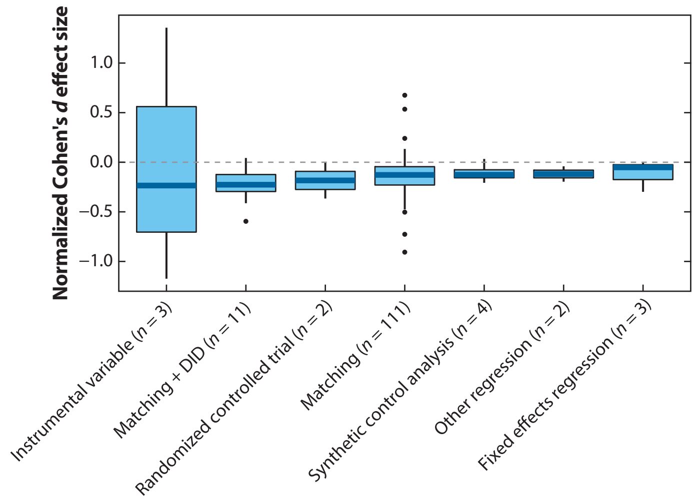

Figure 5

Standardized forest conservation effect sizes (Cohen's *d*) by evaluation method ($n = 136$). Boxes show the interquartile range, horizontal bars show the median, whiskers extend to the smallest/largest value no further than 1.5 times the interquartile range from the box, and the dots are outliers. Negative values indicate avoided deforestation. Abbreviation: DID, difference-in-difference.

and PES are moderately effective, while certification prevented forest loss only marginally. In some cases, the effects of certification were not significantly different from zero (Blackman et al. 2018). For the reasons mentioned above, we caution against interpreting these findings as guidance on which instrument to choose over another. Usually there are valid theoretical arguments to favor, say, a protected area over a PES scheme in a given intervention context (see

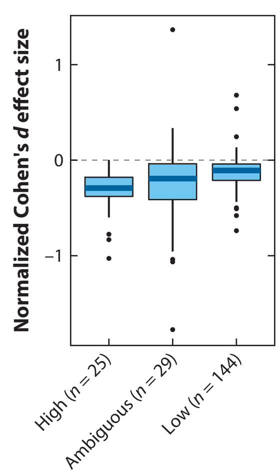

Figure 6

Standardized forest conservation effect sizes (Cohen's *d*) by deforestation pressure ($n = 198$) refers to the sample of Cohen's *d* observations before averaging, so that regional and temporal differences in deforestation pressure could be considered). Boxes show the interquartile range, horizontal bars show the median, whiskers extend to the smallest/largest value no further than 1.5 times the interquartile range from the box, and the dots are outliers. Negative values indicate avoided deforestation.

Section 2). But the outcomes we observe here are as much a function of policy instrument choice as they are mediated by intervention design and unobserved static as well as dynamic context factors.

Effect sizes also vary substantially by choice of evaluation method (**Figure 5**). Given small sample sizes for most methods, these differences most likely also reflect variations in intervention contexts. Instrumental variable analyses exhibited the largest variation in estimated effect sizes, perhaps reflecting the well-known econometric challenges of selecting relevant instruments that fulfill the exclusion restriction. Compared to matching as the most widely used evaluation approach, the effect estimates obtained by fixed effects regressions appear somewhat smaller. Although we would expect that panel models with individual and time effects capture a larger share of the unobservable bias in observational studies than purely matching-based approaches, much larger numbers of independently estimated effects are required to draw robust conclusions on the role of method choices in this field of evaluation research.

We did not find significant differences in effect sizes between countries or continents. This does not imply that intervention contexts, e.g., the pre-existing deforestation pressure, do not matter. Deforestation hotspots tend to be a local phenomenon subject to considerable spatial and temporal variability (Harris et al. 2017). Given the predominance of quasi-experimental evaluation techniques employed in the reviewed literature, we must expect that both program selection criteria and self-selection motives still account for a nontrivial share of the variation among the effect sizes analyzed here. Moreover, even internally valid RCT-based studies may suffer from a location bias introduced by program-targeting criteria.

Based on the qualitative description of the respective local context, we classified all original observations into either high or low deforestation pressure whenever possible. Deforestation pressure was judged using information on intervention placement criteria such as policy objectives and context. For 20% of the observations, mostly including sustainable use areas and decentralized forest management interventions, it was not possible to assign a clear category, so these observations are treated as ambiguous. **Figure 6** shows the magnitude of the location bias after removing the two outlier studies discussed above: Interventions in a high-pressure context have significantly larger (in absolute terms) effect sizes than low-pressure contexts ($p = 0.0002$).

Figure 6 is just a small step toward a standardized comparison of conservation effectiveness across context characteristics, such as strata of human pressure on forest (or natural resources in general). However, our findings here and the conceptual considerations in Sections 1 and 2 lead us to establish the hypothesis that the evaluation literature is thus far biased toward studies that find small, though significant, forest conservation effects due to a high and far location bias of conservation initiatives and a positive publication bias, respectively (Franco et al. 2014, Pirard et al. 2019).

3.4. Development Trade-Offs or Cobenefits of Forest Conservation Policies

Our systematic review above focuses on forest outcomes of conservation policies. A comprehensive review of the impacts of forest conservation policies and initiatives on human well-being other outcomes lies beyond the scope of this paper. Here, we nonetheless briefly summarize the key findings of recent evaluation and review studies with respect to development trade-offs and socioeconomic cobenefits of conservation.

A sizeable literature covers the impacts of PAs on human well-being (Brockington & Wilkie 2015). In a recent global-scale evaluation, Naidoo et al. (2019) find no evidence that proximity to PAs negatively affected a series of alternative well-being. On the contrary, a number of well-being indicators, including household wealth, tended to be positively correlated with proximity to PAs.

In their review of PES impacts on environmental and socioeconomic outcomes in low- and middle-income countries, Snilsveit et al. (2019) find only weak evidence for small positive impacts of PES programs on total household income, but no systematic evidence for negative effects on income or asset related well-being indicators (see also Wunder et al. 2020).

Oya et al. (2018) reviewed socioeconomic impacts of certification systems for agricultural production, including sustainability certification schemes that reward complementary forest conservation outcomes. They find no significant impacts on household income or asset endowments in developing countries.

The comparatively smaller literature rigorously evaluating forest ownership decentralization was synthesized by Samii et al. (2014b). They identified only three studies using quasi-experimental approaches to evaluate well-being outcomes, which found that decentralization initiatives had largely positive effects on forest-specific and general household income measures.

In summary, with respect to the hypothesized strong trade-off between contemporary forest conservation initiatives and human well-being (Roe 2008), evidence from counterfactual-based evaluation studies is still limited for most forest conservation policy instruments. This applies in particular to simultaneous evaluations of socioeconomic outcomes (see Sims & Alix-Garcia 2017), with pronounced knowledge gaps in African countries and for less-intensively studied interventions, such as forest ownership decentralization. Following our ToC (**Figure 1**), small average conservation effects of forest policies and programs can be due to a range of factors, including poor spatial targeting, leakage, or noncompliance. The latter could partially explain the dearth of evaluation studies confirming theoretical conjectures of conservation-poverty trade-offs.

4. CONSERVATION IMPACTS IN THE CONTEXT OF DEVELOPMENT

Our review considers various instruments for forest conservation and multiple methods for impact estimation. As we deliberately apply content and quality filters that significantly reduce the number of studies within a sample usable for Section 3,[1] we also limit our comparison of impacts across relevant contexts. Here, we stress two important, interrelated conservation contexts: economic pressure and development policy.

4.1. Economic Pressure Limits Conservation Impact

Deforestation pressures from private land uses often vary across landscapes. Since J.H. von Thünen, many analysts assert that clearing pressure falls as one moves outward from market centers (Angelsen 2010). Thus, we expect denser forest landscapes farther from markets, given higher transport costs, unless others favor agricultural conversion (e.g., high slopes near markets may stay forested, whereas forests in fertile areas far from markets may be cleared). In evaluation studies, some but not all of these factors can be observed and controlled for.

Variation in pressures across forest landscapes implies heterogeneous impacts of conservation initiatives (Börner et al. 2015). Even if all policies are perfectly enforced, in the absence of pressure (e.g., a PA in remote regions with poor soils and steep slopes) those policies cannot avoid any forest clearing. Hence, impact depends on location. Quasi-experimental evaluation studies support these

[1] Putting this another way, many studies are excluded in Section 3. For instance, there is a vast literature on the impacts of PAs that uses inside-outside comparisons, without checking similarity (e.g., via matching), including when the outside might be defined as all forest outside the PA (versus the perhaps more similar but also more likely spillover-affected lands just outside PAs). That most of the literature finds higher impacts suggests a clear, significant bias in impact estimates from comparing with controls that are not similar. It also implies that PAs would have more impact if created and enforced in higher-pressure areas.

theoretical predictions. Even for a small country like Costa Rica, Pfaff et al. (2009) show significant variations in protection effectiveness along pressure gradients, while Pfaff et al. (2015) demonstrate that huge areas outside of "the arc of deforestation" in the Brazilian Amazon can be sufficiently isolated to have no measurable impacts on deforestation.

Why would the PA be remote? Development trade-offs tend to push conservation initiatives toward low-pressure zones because high opportunity costs, i.e., high agricultural profits, make land expensive to buy for conservation. The individuals interested in those profits will tend to resist the establishment of a PA. Global studies (e.g., Joppa & Pfaff 2009, Nelson & Chomitz 2011) confirm this location tendency.

The net impact of a forest policy includes not only the impacts inside the intervention boundary but also any spillover impacts outside, i.e., policy-induced leakage effects. We must also address those when thinking about heterogeneous impacts. However, leakage is unlikely if there were no significant impacts inside the intervention boundary. Herrera et al. (2019) confirm this for the Brazilian Amazon, while also examining differences between actors for some regions where PA impacts and spillovers are significant.

4.2. Development and Conservation Policy Interactions

It follows from the discussion above that development policies that are successful in stimulating economic activity also affect the impacts of conservation. For instance, investments in roads, railways, or ports and policies facilitating trade and migration significantly affect spatial patterns of economic activity. This not only influences forests directly but also affects where in the landscape conservation has high versus low impacts on forests.

Awareness of the role of development policies for forest conservation and conservation policy impacts is growing (Ibisch et al. 2016, Peinhardt et al. 2019). For instance, results discussed above for heterogeneous PA impacts imply that when new highways cut across frontiers, pressure is eventually likely to arrive and then PAs could importantly constrain deforestation. Such a scenario also raises the possibility of integrating the two kinds of policies, e.g., building a road to link economic hubs, but placing PAs alongside the road's path.

Another way to integrate policy planning is to adjust where and how development policies are implemented. Roads have greatly shaped economic activity and thereby are associated with an enormous share of deforestation over time. Yet, for a goal of linking cities, for instance, more than one possible route often exists. Similarly, energy pipelines can be sited and implemented to minimize forest loss and more generally aid conservation, as development policies' impacts vary with context.

Adjusting development policy could, effectively, be considered conservation policy. Pfaff et al. (2018), for example, showed that new roads might significantly shift land use and deforestation on the frontier when conditions are poised to become favorable for economic development. In contrast, road improvements in places with significant prior road investment tend to cause lower additional forest loss. At the other extreme, if a new road connects distant isolated frontiers, without a flexible supply of labor and capital, the short-run impacts may be smaller. Hence, the interaction between development and conservation is scenario specific (Angelsen & Kaimowitz 2001). Going beyond the frequently studied context of Latin American frontiers, much of the world features far higher prior economic development. Following the logic above, that should lower average forest impacts from new roads, as has been found by Kaczan (2020) for new roads due to an investment program in India.

Yet, sometimes development policies instead lead to the elimination of conservation. PA downgrading, downsizing, and degazettement (PADDD) events are forms of reducing protection typically associated with both infrastructure investments and other land-use pressures. Their impacts can be reduced via effective bargaining, e.g., surrendering less-effective PAs (see Tesfaw et al. 2018). Conversely, negotiating and enforcing conservation can affect development policy impacts. For instance, Herrera (2015) shows that protected area siting can signal that the region is not going to be a center of pro-development investment. This can induce migration and, due to the public response to both the PAs and the migration, shift investments in either maintaining or creating new roads.

5. CONCLUSIONS AND OUTLOOK

This review expands the scope of earlier reviews of the effectiveness of forest conservation policies, including Börner et al. (2016), and provides complementary evidence to contemporary review studies, such as Burivalova et al. (2019). We document that theoretical thinking about the determinants of forest conservation effectiveness has advanced to include complex interactions of mediating contextual factors and policy design features that determine both environmental outcomes and socioeconomic cobenefits of forest conservation policies (Section 2). Taking these lessons on board is a challenge not only for the design of effective policies, including coordination across policy sectors, but also for research that aims at closing the numerous gaps of evidence in the literature evaluating forest conservation initiatives. The possible combinations between contextual factors, implementation modes, and policy mixes are so manifold that safely attributing impact evaluation studies would appear an impossible mission: There may be too much variation out there to statistically control for. By necessity, we are thus rather forced to patch together an inevitably incomplete and yet over time increasingly concrete picture of what type of conservation tools may work under what circumstances.

Our review rests on a database of 167 full texts that were identified using a documented set of filtering criteria and human expert screening supported by machine learning. We construct a sample of 136 normalized effect sizes for 10 categories of conservation tools. Biases in our sample of studies may result from filtering criteria and the location biases of both policy initiatives and their evaluators as well as the presence (or not) of basic descriptive statistics in the respective peer-reviewed literature. Comparisons of effect sizes across instruments, evaluation methods, and intervention contexts, nonetheless, provide interesting insights that are largely consistent with theoretical expectations (Section 2).

First, we find that forest conservation policies exhibit average effects in the lower Cohen's d range (around 0.2), which are comparable to effect size ranges in the development sector. From the evidence reviewed here, no forest conservation policy instrument qualifies a priori as a first best choice for policy makers, including because our sample is too small to allow for robust comparison across instruments. So far, nonetheless, the protection of indigenous lands ($n = 3$) and incentive-based conservation tools, such as PES ($n = 22$) and ICDP ($n = 3$), range at the higher end of the effectiveness spectrum (see **Figure 4**). A caveat here is that we know little about the cost-efficiency of these interventions.

Second, we make further progress toward confirming the importance of intervention contexts, particularly preintervention pressure on forests, in determining forest outcomes. Most of the studies in our sample evaluated conservation initiatives in comparatively low-pressure contexts and thus exhibited significantly lower conservation impacts than studies focusing on interventions in high-pressure settings. Our extended discussion of determinants of heterogeneous conservation impacts correspondingly concentrates on the importance of location for impact. This discussion further emphasizes that development policies are important drivers of deforestation and mediators of conservation policy impacts and thus may need to be redesigned and coordinated to align with forest conservation objectives.

We note that the success of many biodiversity or environmental service-focused interventions may be proxied fairly well by forest cover change, our comparative indicator used above. Sometimes this proxy may be deceptive, though. The more spatially specific a conservation target (e.g., threatened species protection, erosion protection in a fragile watershed, roadside recreational benefits), the more ambiguous the relationship becomes: In principle, interventions could overall have protected little forest, but exactly the right forest in terms of environmental service provision. Likewise, in higher- to middle-income countries with forest regrowth, environmental gains correlate less well with forest cover increments (e.g., fire protection, biodiversity, or recreation in forest-farm fragments).

Third, we compare conservation effect sizes across evaluation methods and find sizable variance across and within evaluation method categories. We suspect that this variance is driven largely by variations in intervention contexts as opposed to violations of internal validity—a common critique of the quasi-experimental methods that dominate in this field (Jayachandran et al. 2017). Clearly, much larger numbers of studies using both RCT and quasi-experimental methods are needed to confirm this conjecture. However, given the peculiarities of the conservation sector (see Section 1), including the dominance of disincentive-based interventions in nature conservation, we would expect quasi-experimental methods to dominate the field for some time.

This and other reviews cited above illustrate that impact evaluation has developed into a dynamic new field of research in the conservation sector. Waiting for more and better evidence can, however, not serve as an excuse for business as usual in conservation policy design. The early evaluation-based evidence tends to confirm key lessons from the large body of theoretical and case-based empirical research that should inform policy design and cross-sectoral coordination for forest conservation already today (Seymour & Harris 2019, Wunder et al. 2018). Further progress toward evidence-based forest conservation requires (*a*) a more explicit consideration of policy and program implementation costs to enrich evaluation research with measures of cost-efficiency, (*b*) a more systematic approach to evaluating interventions across context strata of deforestation and economic pressure, (*c*) more frequent consideration of joint environmental and social outcomes and conservation schemes, and (*d*) additional efforts to evaluate conservation outcomes of development and other relevant nonforest sector interventions.

DISCLOSURE STATEMENT

The authors are not aware of any affiliations, memberships, funding, or financial holdings that might be perceived as affecting the objectivity of this review.

ACKNOWLEDGMENTS

This research has been supported by the German Federal Ministry of Education and Research (BMBF) as part of the project STRIVE (Sustainable Trade and Innovation Transfer in the Bioeconomy) and by NORAD's Global Comparative Study on REDD+. J.B. acknowledges funding by the German Research Foundation (DFG) through the collaborative research center "Future Rural Africa," funding code TRR 228/1.

LITERATURE CITED

Abman R. 2018. Rule of law and avoided deforestation from protected areas. *Ecol. Econ.* 146:282–89

Alix-Garcia JM, Sims KRE, Orozco-Olvera VH, Costica LE, Fernández Medina JD, Monroy SR. 2018. Payments for environmental services supported social capital while increasing land management. *PNAS* 115(27):7016–21

Ambec S, Cohen MA, Elgie S, Lanoie P. 2013. The Porter hypothesis at 20: Can environmental regulation enhance innovation and competitiveness? *Rev. Environ. Econ. Policy* 7(1):2–22

Andrews RNL. 2006. *Managing the Environment, Managing Ourselves: A History of American Environmental Policy*. New Haven, CT: Yale Univ. Press. 2nd ed.

Angelsen A. 2010. Policies for reduced deforestation and their impact on agricultural production. *PNAS* 107(46):19639–44

Angelsen A, Kaimowitz D, eds. 2001. *Agricultural Technologies and Tropical Deforestation*. Oxon, UK: CABI Int.

Arima EY, Barreto P, Araújo E, Soares-Filho B. 2014. Public policies can reduce tropical deforestation: lessons and challenges from Brazil. *Land Use Policy* 41:465–73

Arriagada RA, Sills EO, Pattanayak SK. 2011. *Payments for environmental services and their impact on forest transition in Costa Rica*. Work. Pap., Dep. Agric. Econ., Univ. Cátol., Santiago

Baylis K, Honey-Rosés J, Börner J, Corbera E, Ezzine-de-Blas D, Ferraro PJ, et al. 2016. Mainstreaming impact evaluation in nature conservation. *Conserv. Lett.* 9(1):58–64

Blackman A, Corral L, Lima ES, Asner GP. 2017. Titling indigenous communities protects forests in the Peruvian Amazon. *PNAS* 114(16):4123–28

Blackman A, Goff L, Rivera Planter M. 2018. Does eco-certification stem tropical deforestation? Forest Stewardship Council certification in Mexico. *J. Environ. Econ. Manag.* 89:306–33

Börner J, Baylis K, Corbera E, Ezzine-de-Blas D, Ferraro PJ, Honey-Rosés J, et al. 2016. Emerging evidence on the effectiveness of tropical forest conservation. *PLOS ONE* 11(11):e0159152

Börner J, Marinho E, Wunder S. 2015. Mixing carrots and sticks to conserve forests in the Brazilian Amazon: a spatial probabilistic modeling approach. *PLOS ONE* 10(2):e0116846

Börner J, Vosti SA. 2013. Managing tropical forest ecosystem services: an overview of options. In *Governing the Provision of Ecosystem Services*, Vol. 4, ed. R Muradian, L Rival, pp. 21–46. Dordrecht: Springer

Börner J, Wunder S, Reimer F, Bakkegaard RK, Viana V, et al. 2013. *Promoting Forest Stewardship in the Bolsa Floresta Programme: Local Livelihood Strategies and Preliminary Impacts*. Rio de Janeiro: CIFOR. **https://www.cifor.org/library/4481/**

Bowler DE, Buyung-Ali LM, Healey JR, Jones JPG, Knight TM, Pullin AS. 2012. Does community forest management provide global environmental benefits and improve local welfare? *Front. Ecol. Environ.* 10(1):29–36

Brockington D, Wilkie D. 2015. Protected areas and poverty. *Philos. Trans. R. Soc. B* 370(1681):20140271

Burivalova Z, Miteva D, Salafsky N, Butler RA, Wilcove DS. 2019. Evidence types and trends in tropical forest conservation literature. *Trends Ecol. Evol.* 34(7):669–79

Caplow S, Jagger P, Lawlor K, Sills E. 2011. Evaluating land use and livelihood impacts of early forest carbon projects: lessons for learning about REDD+. *Environ. Sci. Policy* 14(2):152–67

Cashore B, Gale F, Meidinger E, Newsom D. 2006. Forest certification in developing and transitioning countries: Part of a sustainable future? *Environ. Sci. Policy Sustain. Dev.* 48(9):6–25

Coad L, Leverington F, Knights K, Geldmann J, Eassom A, et al. 2015. Measuring impact of protected area management interventions: current and future use of the Global Database of Protected Area Management Effectiveness. *Philos. Trans. R. Soc. B* 370(1681). **https://doi.org/10.1098/rstb.2014.0281**

De Freitas FLM, Sparovek G, Mörtberg U, Silveira S, Klug I, Berndes G. 2017. Offsetting legal deficits of native vegetation among Brazilian landholders: effects on nature protection and socioeconomic development. *Land Use Policy* 68:189–99

De Sy V, Herold M, Achard F, Avitabile V, Baccini A, et al. 2019. Tropical deforestation drivers and associated carbon emission factors derived from remote sensing data. *Environ. Res. Lett.* 14(9):94022

Engel S. 2016. The devil in the detail: a practical guide on designing payments for environmental services. *Int. Rev. Environ. Resour. Econ.* 9(1–2):131–77

Ferraro PJ. 2009. Counterfactual thinking and impact evaluation in environmental policy. *New Dir. Eval.* 2009(122):75–84

Ferraro PJ, Kiss A. 2002. Direct payments to conserve biodiversity. *Science* 298(5599):1718–19

Franco A, Malhotra N, Simonovits G. 2014. Publication bias in the social sciences: unlocking the file drawer. *Science* 345(6203):1502–5

Geldmann J, Barnes M, Coad L, Craigie ID, Hockings M, Burgess ND. 2013. Effectiveness of terrestrial protected areas in reducing habitat loss and population declines. *Biol. Conserv.* 161:230–38

Harahap F, Silveira S, Khatiwada D. 2017. Land allocation to meet sectoral goals in Indonesia—an analysis of policy coherence. *Land Use Policy* 61:451–65

Harris NL, Goldman E, Gabris C, Nordling J, Minnemeyer S, Ansari S, et al. 2017. Using spatial statistics to identify emerging hot spots of forest loss. *Environ. Res. Lett.* 12(2):024012

Heltberg R. 2001. Determinants and impact of local institutions for common resource management. *Envir. Dev. Econ.* 6(2):183–208

Herrera D. 2015. *Protected areas' deforestation spillovers and two critical underlying mechanisms: an empirical exploration for the Brazilian Amazon*. PhD Thesis, Duke Univ., Durham, NC

Herrera D, Pfaff A, Robalino J. 2019. Impacts of protected areas vary with the level of government: comparing avoided deforestation across agencies in the Brazilian Amazon. *PNAS* 116(30):14916–25

Ibisch PL, Hoffmann MT, Kreft S, Pe'er G, Kati V, et al. 2016. A global map of roadless areas and their conservation status. *Science* 354(6318):1423–27

Jayachandran S, De Laat J, Lambin EF, Stanton CY, Audy R, Thomas NE. 2017. Cash for carbon: a randomized trial of payments for ecosystem services to reduce deforestation. *Science* 357(6348):267–73

Joppa LN, Pfaff A. 2009. High and far: biases in the location of protected areas. *PLOS ONE* 4(12):e8273

Kaczan DJ. 2020. Can roads contribute to forest transitions? *World. Dev.* 129:104898

Kaimowitz D, Angelsen A. 1998. *Economic Models of Tropical Deforestation: A Review*. Bogor, Indones.: CIFOR

Lambin EF, Gibbs HK, Heilmayr R, Carlson KM, Fleck LC, et al. 2018. The role of supply-chain initiatives in reducing deforestation. *Nat. Clim. Change* 8(2):109–16

Lambin EF, Turner BL, Geist HJ, Agbola SB, Angelsen A, et al. 2001. The causes of land-use and land-cover change: moving beyond the myths. *Glob. Environ. Change* 11(4):261–69

Leisher C, Temsah G, Booker F, Day M, Samberg L, et al. 2016. Does the gender composition of forest and fishery management groups affect resource governance and conservation outcomes? A systematic map. *Environ. Evid.* 5:6

Liscow ZD. 2013. Do property rights promote investment but cause deforestation? Quasi-experimental evidence from Nicaragua. *J. Environ. Econ. Manag.* 65(2):241–61

Macura B, Secco L, Pullin AS. 2015. What evidence exists on the impact of governance type on the conservation effectiveness of forest protected areas? Knowledge base and evidence gaps. *Environ. Evid.* 4:24

Mahanty S, McDermott CL. 2013. How does 'Free, Prior and Informed Consent'(FPIC) impact social equity? Lessons from mining and forestry and their implications for REDD+. *Land Use Policy* 35:406–16

Manning CD, Raghavan P, Schütze H. 2008. *Introduction to Information Retrieval*. Cambridge, UK: Cambridge Univ. Press. 4th ed.

Naidoo R, Gerkey D, Hole D, Pfaff A, Ellis AM, et al. 2019. Evaluating the impacts of protected areas on human well-being across the developing world. *Sci. Adv.* 5(4):eaav3006

Nelson A, 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(8):e22722

Nkonya E, Mirzabaev A, von Braun J, eds. 2016. *Economics of Land Degradation and Improvement—A Global Assessment for Sustainable Development*. Cham, Switz.: Springer Open

Ojanen M, Zhou W, Miller DC, Nieto SH, Mshale B, Petrokofsky G. 2017. What are the environmental impacts of property rights regimes in forests, fisheries and rangelands? *Environ. Evid.* 6:12

Oldekop JA, Holmes G, Harris WE, Evans KL. 2016. A global assessment of the social and conservation outcomes of protected areas. *Conserv. Biol.* 30(1):133–41

Oya C, Schaefer F, Skalidou D. 2018. The effectiveness of agricultural certification in developing countries: a systematic review. *World Dev*. 112:282–312

Paiva RJO, Brites RS, Machado RB. 2015. The role of protected areas in the avoidance of anthropogenic conversion in a high pressure region: a matching method analysis in the core region of the Brazilian Cerrado. *PLOS ONE* 10(7):e0132582

Parks T. 2008. The rise and fall of donor funding for advocacy NGOs: understanding the impact. *Dev. Pract.* 18(2):213–22

Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, et al. 2011. Scikit-learn: machine learning in Python. *J. Mach. Learn. Res.* 12:2825–30

Peinhardt C, Kim AA, Pavon-Harr V. 2019. Deforestation and the United States-Peru Trade Promotion Agreement. *Glob. Environ. Politics* 19(1):53–76

Petrosino A, Morgan C, Fronius TA, Tanner-Smith EE, Boruch RF. 2012. Interventions in developing nations for improving primary and secondary school enrollment of children: a systematic review. *Campbell Syst. Rev.* 8(1). **https://doi.org/10.4073/csr.2012.19**

Pfaff A, Robalino J. 2017. Spillovers from conservation programs. *Annu. Rev. Resour. Econ.* 9:299–315

Pfaff A, Robalino J, Herrera D, Sandoval C. 2015. Protected areas' impacts on Brazilian Amazon deforestation: examining conservation-development interactions to inform planning. *PLOS ONE* 10(7):e0129460

Pfaff A, Robalino J, Sánchez-Azofeifa GA, Andam KS, Ferraro PJ. 2009. Park location affects forest protection: land characteristics cause differences in park impacts across Costa Rica. *BE J. Econ. Anal. Policy* 9(2). **https://doi.org/10.2202/1935-1682.1990**

Pfaff ASP, Robalino J, Reis EJ, Walker R, Perz S, Laurance W, et al. 2018. Roads & SDGs, tradeoffs and synergies: learning from Brazil's Amazon in distinguishing frontiers. *Econ. Open-Access Open-Assess. E-J.* 12(2018–11):1–26

Pirard R, Wunder S, Duchelle AE, Puri J, Asfaw S, et al. 2019. *Effectiveness of forest conservation interventions: an evidence gap map*. IEU Learn. Pap. 02, Green Climate Fund Indep. Eval. Unit, Incheon, Rep. Korea. **https://efi.int/sites/default/files/files/efimed/resources/Pirard%20et%20al.%20EGM%20Cons%20IEU%202019.pdf**

Porter-Bolland L, Ellis EA, Guariguata MR, Ruiz-Mallén I, Negrete-Yankelevich S, Reyes-García V. 2012. Community managed forests and forest protected areas: an assessment of their conservation effectiveness across the tropics. *Forest Ecol. Manag.* 268:6–17

Rana P, Sills E. 2018. Does certification change the trajectory of tree cover in working forests in the tropics? An application of the synthetic control method of impact evaluation. *Forests* 9(3):98

Robinson EJZ, Kumar AM, Albers HJ. 2010. Protecting developing countries' forests: enforcement in theory and practice. *J. Nat. Resour. Policy Res.* 2(1):25–38

Roe D. 2008. The origins and evolution of the conservation-poverty debate: a review of key literature, events and policy processes. *Oryx* 42(4):491–503

Rudel TK. 2017. The dynamics of deforestation in the wet and dry tropics: a comparison with policy implications. *Forests* 8(4):108

Samii C, Lisiecki M, Kulkarni P, Paler L, Chavis L. 2014a. Effects of decentralized forest management (DFM) on deforestation and poverty in low and middle income countries: a systematic review. *Campbell Syst. Rev.* 10. **https://doi.org/10.4073/csr.2014.10**

Samii C, Lisiecki M, Kulkarni P, Paler L, Chavis L. 2014b. Effects of payment for environmental services (PES) on deforestation and poverty in low and middle income countries: a systematic review. *Campbell Syst. Rev.* 10. **https://doi.org/10.4073/csr.2014.11**

Sans GHC, Aguiar S, Vallejos M, Paruelo JM. 2018. Assessing the effectiveness of a land zoning policy in the Dry Chaco. The case of Santiago del Estero, Argentina. *Land Use Policy* 70:313–21

Sawilowsky SS. 2009. New effect size rules of thumb. *J. Mod. Appl. Stat. Methods* 8(2):26

Scholes R, Montanarella L, Brainich E, Barger N, ten Brink B, et al. 2018. *Summary for policymakers of the assessment report on land degradation and restoration of the Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services*. Rep., IPBES Secr., Bonn, Ger.

Sen B. 1987. NGO self-evaluation: issues of concern. *World Dev.* 15:161–67

Seymour F, Harris NL. 2019. Reducing tropical deforestation. *Science* 365(6455):756–57

Sills EO, Herrera D, Kirkpatrick AJ, Brandão A, Dickson R, et al. 2015. Estimating the impacts of local policy innovation: the synthetic control method applied to tropical deforestation. *PLOS ONE* 10(7):e0132590

Sims KRE, Alix-Garcia JM. 2017. Parks versus PES: Evaluating direct and incentive-based land conservation in Mexico. *J. Environ. Econ. Manag.* 86:8–28

Snilsveit B, Stevenson J, Langer L, Polanin J, Shemilt I, et al. 2019. Incentives for climate mitigation in the land use sector—the effects of payment for environmental services on environmental and socioeconomic outcomes in low-and middle-income countries: a mixed-methods systematic review. *Campbell Syst. Rev.* 15(3):e1045

Sunderlin WD, Hatcher J, Liddle M. 2008. *From Exclusion to Ownership? Challenges and Opportunities in Advancing Forest Tenure Reform*. Washington, DC: Rights Resour. Init. **https://www.cifor.org/library/5869/**

Tesfaw AT, Pfaff A, Kroner REG, Qin S, Medeiros R, Mascia MB. 2018. Land-use and land-cover change shape the sustainability and impacts of protected areas. *PNAS* 115(9):2084–89

Tilbury D. 1995. Environmental education for sustainability: defining the new focus of environmental education in the 1990s. *Environ. Educ. Res.* 1(2):195–212

Vatn A. 2005. Rationality, institutions and environmental policy. *Ecol. Econ.* 55(2):203–17

Vlek PLG, Khamzina A, Tamene L, eds. 2017. *Land degradation and the Sustainable Development Goals: threats and potential remedies*. CIAT Publ. 440, Int. Cent. Trop. Agric., Nairobi

Wahlén CB. 2014. Constructing conservation impact: understanding monitoring and evaluation in conservation NGOs. *Conserv. Soc.* 12(1):77–88

Weiss CH. 1997. How can theory-based evaluation make greater headway? *Eval. Rev.* 21(4):501–24

Wiik ER, d'Annunzio R, Pynegar E, Crespo D, Asquith N, Jones JPG. 2019. Experimental evaluation of the impact of a payment for environmental services program on deforestation. *Conserv. Sci. Pract.* 1(2):e8

Wittemyer G, Elsen P, Bean WT, Burton ACO, Brashares JS. 2008. Accelerated human population growth at protected area edges. *Science* 321(5885):123–26

Wunder S. 2015. Revisiting the concept of payments for environmental services. *Ecol. Econ.* 117:234–43

Wunder S, Börner J, Ezzine-de-Blas D, Feder S, Pagiola S. 2020. Payments for environmental services: past performance and pending potentials. *Annu. Rev. Resour. Econ.* 12. In press

Wunder S, Brouwer R, Engel S, Ezzine-de-Blas D, Muradian R, Pascual U, Pinto R. 2018. From principles to practice in paying for nature's services. *Nat. Sustain.* 1(3):145–50