

Cyrus Samii
Matthew Lisiecki
Parashar Kulkarni
Laura Paler
Larry Chavis

Payment for environmental services for reducing deforestation and poverty in low- and middle-income countries

A systematic review

March 2015

Systematic
Review 17

Environment



International
Initiative for
Impact Evaluation

About 3ie

The International Initiative for Impact Evaluation (3ie) is an international grant-making NGO promoting evidence-informed development policies and programmes. We are the global leader in funding, producing and synthesising high-quality evidence of what works, for whom, why and at what cost. We believe that better and policy-relevant evidence will make development more effective and improve people's lives.

3ie systematic reviews

3ie systematic reviews appraise and synthesise the available high-quality evidence on the effectiveness of social and economic development interventions in low- and middle-income countries. These reviews follow scientifically recognised review methods, and are peer-reviewed and quality assured according to internationally accepted standards. 3ie is providing leadership in demonstrating rigorous and innovative review methodologies, such as using theory-based approaches suited to inform policy and programming in the dynamic contexts and challenges of low- and middle-income countries.

About this review

Payment for environmental services for reducing deforestation and poverty in low- and middle-income countries: a systematic review, was submitted in partial fulfilment of the requirements of SR4.1173 issued under Systematic Review Window 4. This review is available on the [3ie website](#). 3ie is publishing this report as received from the authors; it has been formatted to 3ie style. This review has also been published in the Campbell Collaboration Library and is available [here](#).

All content is the sole responsibility of the authors and does not represent the opinions of 3ie, its donors or its board of commissioners. Any errors are also the sole responsibility of the authors. Comments or queries should be directed to the corresponding author, Cyrus Samii at [cgs2083@nyu.edu](mailto:cds2083@nyu.edu)

Funding for this systematic review was provided by the Norwegian Agency for Development Cooperation.

Suggested citation: 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).

3ie systematic review executive editors: Philip Davies and Beryl Leach

Managing editor: Deepthy Menon

Technical editor: Birte Snilstveit

Production manager: Pradeep Singh

Cover design: John F McGill and Akarsh Gupta

© International Initiative for Impact Evaluation (3ie), 2015, 2016

Payment for environmental services for reducing deforestation and poverty in low- and middle-income countries: a systematic review

Cyrus Samii
Department of Politics, New York University

Matthew Lisiecki
Department of Politics, New York University

Parashar Kulkarni
Department of Politics, New York University

Laura Paler
Department of Political Science, University of Pittsburgh

Larry Chavis
Kenan-Flagler Business School, University of North Carolina

3ie Systematic Review 17

March 2015



Abstract

We conducted a systematic review of studies on the impact of payments for environmental services (PES) that set natural forest conservation as the goal on deforestation and poverty in developing countries. The review is motivated by debates over whether the pursuits of conservation and poverty reduction in developing countries tend to conflict or whether they might be complementary. A search for rigorous impact evaluation studies identified eleven quantitative and nine associated qualitative evaluation studies assessing the effects of PES. The methodological rigor of these studies varied widely, meaning that the evidence base for the impact of PES policies is limited in both quantity and quality. Given the evidence available, we find little reason for optimism about the potential for current PES approaches to achieve both conservation and poverty reduction benefits jointly. We call for the production of high quality impact evaluations, using randomisation when possible, to assess whether the apparent incompatibility of conservation and poverty reduction might be overcome through programming innovations.

Summary

Background

Natural forest preservation in the tropics, and thus in developing countries, must be an element of any effective effort to manage climate change. Forests serve as natural carbon sinks, which help to mitigate the effect of other carbon emissions. However, forest cover is being reduced and it is estimated that deforestation is responsible for 10-17 per cent of global carbon emissions. Since 2007, governments have coordinated conservation efforts under the Reducing Emissions through Deforestation and Forest Degradation (REDD+) initiative, which has led to the implementation of various programs designed to reduce the amount of forested land converted to other purposes.

Payment for environmental services (PES) programs is one type of intervention commonly implemented under the REDD+ umbrella. PES programs allow for direct exchange between those demanding 'environmental services' such as protection or rehabilitation of natural forests and those in a position to provide them locally. While the primary goal of reducing deforestation is clear, the policy and academic literature debates the extent to which PES programs in developing countries should incorporate goals of poverty reduction. Some argue that the targeting of poverty goals will undermine conservation effectiveness (e.g., because behavioural change among poorer households does not have as much potential to promote conservation as that of wealthier households or commercial entities). Others argue that targeting benefits toward the poor would contribute to conservation effectiveness by either promoting sustainable livelihoods or helping to legitimize conservation programming.

In this review, we assess the effects of PES programs on deforestation and welfare outcomes in low and middle income countries (LMICs), and whether the twin goals of improving both environmental and human welfare outcomes are at odds with each other. We also examine how inequality, institutional capacity, corruption, and democratic accountability may moderate the effects of PES programs. Conducting this review is important for moving the debate around PES beyond theoretical discussions and into a better-informed, evidence-based discussion.

Objectives

The first objective of this review is to assess the evidence on the effects of PES interventions on conservation and poverty outcomes in LMICs. A second objective is to assess the extent to which effects on poverty in turn affect whether conservation benefits are realized. The third objective is to evaluate how institutional and social conditions (namely, inequality, institutional capacity, corruption, and democratic accountability) moderate the effects of PES programs.

Selection criteria

The review includes studies of PES programs that assess effects on either (i) deforestation outcomes in forest areas in developing countries or (ii) poverty conditions of populations residing in communities that are proximate to natural growth forest areas in developing countries. We included studies using a range of measures for both deforestation (on-the-ground point samples, samples created from satellite imagery) and welfare (consumption, income, or income potential).

We required that PES programs have a clear start date when either payments or rewards are themselves offered to individual or corporate property holders to maintain or rehabilitate (for example, via planting endemic species) natural forests, or institutions are established to facilitate such offers.

For quantitative synthesis we included (a) randomized studies and (b) quasi-experimental studies that employ strategies for causal identification with clearly delineate treated and control areas and use some method for removing biases due to non-random assignment of the intervention. Qualitative data are used in the synthesis to provide descriptions and context for interventions that are included in the quantitative synthesis. Such data were drawn from the quantitative studies themselves as well as qualitative studies that cover the same programs or settings as the quantitative studies.

Search strategy

To find the articles included in this review, we searched a variety of databases using key words related to PES programs. The set of databases and list of keywords were assembled based on consultation with a Campbell Collaboration information retrieval specialist. We also carried out hand searches of key journals in relevant fields, using publisher search engines and references cited in papers accepted for review as well as in review papers or thematically relevant papers identified during the search.

Data collection and analysis

For studies eligible for inclusion in the review we systematically collected data on study characteristics, findings, and moderators. Risk of bias was assessed based on the guidance of the IDCG Risk of Bias Tool (version March 2012). We extracted qualitative information from both the included quantitative studies as well as qualitative studies that covered the same types of programs and contexts as our quantitative studies. We use such qualitative data to establish that conditions recorded in quantitative data are being interpreted correctly and to provide descriptions and context for interventions that are included in the quantitative synthesis.

For effects on forest cover, whenever possible we standardized them to annual forest cover change rates. For effects on material welfare and poverty outcomes, we used percentage change over estimated average counterfactual outcome (e.g. for income effects, per cent change in income relative to the average income of the control group). For each hypothesis, we synthesised estimates using meta-analysis when the following conditions were met: (i) more than two studies meeting the quantitative inclusion criteria; (ii) effect sizes for common outcome constructs; and (iii) effects measured against similar comparators.

Results

Our database search returned 1382 articles on PES programs. After eliminating articles that were not relevant to our hypotheses or conducted with appropriate methodological rigor, we were left with 20 articles on PES programs. Of these 11 PES articles conducted quantitative impact evaluation of these programs. The 11 PES articles cover six programs in four countries (Costa Rica, China, Mexico, and Mozambique).

The resulting evidence base is weak both in terms of the number of eligible studies and the methodological weaknesses of the included studies. None of the studies are based on

randomized experiments, and so the potential for hidden selection or confounding biases is the most concerning issue. Few of the studies create comparison groups that allow them to address spill-over and leakage of effects from program areas to non-program areas. None of the studies investigated forest conservation and welfare effects jointly, which made it difficult to assess how these two goals are related.

Effects on Deforestation Outcomes

The PES studies that assessed programs' effects on forest cover included nine studies of four programs in Costa Rica and Mexico. The studies focused on two types of measures: impact on deforestation rate (where the best-case scenario is a deforestation rate of 0) and impact on forest cover (which allow for a positive outcome in the expansion of forested land). Keeping in mind the weakness of the evidence base, the studies that focus solely on reducing deforestation suggest that PES programs have, on average, tended to reduce the annual rate of deforestation by 0.21 percentage points (s.e.=0.09, 95% CI: [0.03, 0.39]). Effect sizes are a bit larger for studies that examine forest cover change, which measures not only forest loss, but also forest gain. Estimated effects on annual forest cover change rates ranged from 0.50 percentage points (s.e.=0.20, 95% CI: [0.11, 0.89]) for a study in Costa Rica to 1.6 percentage points (s.e.=0.80, 95% CI: [0.03, 3.17]) for a study in Mexico. One study suggests an outlier effect of 10 percentage points (no standard error or confidence interval provided in the original study), but this study suffers from a high risk of bias.

Effects on Human Welfare Outcomes

The evidence base on the effects of PES on welfare outcomes is very limited, with only two studies in two countries (China and Mozambique) included in the review. These studies find that PES improved participating households' incomes by 4 per cent (s.e.=1.55, 95% CI: [0.96, 7.04]) in Mozambique and by 14 per cent (s.e.=3.42, 95% CI: [7.3, 20.7]) in China. However, these average effects do not necessarily tell us how these programs affect poor households. For PES to contribute to poverty reduction, poorer households must be able to participate at high rates. But participation in PES programs is typically more difficult for poor households than wealthier households (a fact documented by in a number of the studies included in the review). The study from Mozambique includes estimates for poor households and finds that the welfare effects were substantially less in absolute terms, and not statistically significant for these households.

The Role of Institutional and Social Conditions

We aimed to address a number of hypotheses regarding the influence of institutional and social conditions (inequality, institutional capacity, corruption, and democratic accountability) on the effects of PES programs. However, due to limitations of the evidence base we were unable to test these hypotheses. We did however extract qualitative data from included studies and associated qualitative studies that provide some insights into the role of institutional and social conditions in the context of PES programs.

A study on the Mexican PSAH PES program found that forest conservation effects were worse in poorer areas. Qualitative information from Costa Rica was consistent with this account. Several of the studies also addressed the issue of institutional capacity, describing situations where PES programs did not have the ability to carry out their mandates.

Corruption and possible misappropriation of project resources were also factors raised in a qualitative study of PES in Mexico. The study found that program resources were applied to address inadequacies in other government programs.

Authors' conclusions

Limitations in the evidence base preclude definitive hypothesis tests, however the evidence we find suggests that PES does reduce deforestation rates. The effect is modest however and seems to come with high levels of inefficiency. In terms of PES effects on poverty, we cannot say that the evidence indicates beneficial effects. Available evidence shows that PES programs are less effective in poor areas and are less likely to attract participation of poor households than wealthier ones. These are troubling findings but they are based on only a handful of cases and therefore deserve much more empirical attention.

Our review aimed to assess the extent to which environmental and poverty reduction goals conflict with one another, how different conservation strategies fare in terms of such trade-offs, and the scope for 'win-win' strategies that generate both significant environmental and poverty reduction benefits. Based on the evidence available, we do not find that a case can be made for conservation and poverty-reduction goals being complementary in PES programming.

Our final conclusion re-emphasizes the poor state of the evidence base for PES programming. Much advanced scientific effort and extensive investment has gone into measuring forest conditions around the world. Relative to that, efforts to assess the effects of PES programs on deforestation and poverty are limited and methodologically weak. Researchers should consider the recent work in development economics for guidance on executing field experiments that might provide more credible evidence (Banerjee and Duflo, 2011; Casey et al., 2012; Karlan and Appel, 2012).

Contents

Abstract	1
Summary	ii
List of figures and tables	v
1. Background	1
1.1. Description of the problem	1
1.2. Description of the intervention.....	2
1.3. How the intervention might work	3
1.4. Why it was important to do this review	7
2. Objectives	8
3. Selection Criteria	9
3.1. Participants.....	9
3.2. Interventions	9
3.3. Outcomes	9
3.4. Study Types.....	10
4. Search Strategy	12
4.1. Electronic searches	12
4.2. Other Searches	12
5. Data Collection and Analysis	13
5.1. Data collection and analysis	13
5.2. Data synthesis	15
6. Results	17
6.1. Characteristics of included studies.....	19
6.2. Characteristics of included studies.....	21
6.3. Effects of PES on poverty	25
6.4. Intersection of poverty and deforestation impact.....	28
6.5. ROLE of institutional and social context.....	29
6.6. Variation in the Program Design Features	31
7. Authors' Conclusions	32
7.1. Summary of findings with respect to our hypotheses	32
7.2. Implications for policy: Elusive win-win	32
7.3. Implications for research: Need for more rigorous research across contexts	34
7.4. Limitations and deviations from protocol.....	35
7.5. Acknowledgements	35
References	36
Appendices	46

List of figures and tables

Figure 1: Illustrating the theory of change	4
Figure 2: Study search process.....	17
Figure 3: Risk of bias assessment.....	21
Figure 4: Estimates of the effect of PES on forest cover change rates due to deforestation (top) and due to either deforestation or forest expansion (bottom).....	25
Table 1: Questions and types of evidence needed for the review.....	8
Table 2: PICOS inclusion criteria	11
Table 3: Quantitative PES study characteristics (grouped by program).....	18
Table 4: Qualitative PES study characteristics	19
Table 5: Studies on PES and forest conservation outcomes	24
Table 6: Studies on PES and poverty outcomes	28
Table 7: Program design features for PES programs	31

Appendices figure

Figure A. 1: Forest cover change under the compound interest law.....	52
Figure A. 2: Implications of estimated effects of PES for forest cover change. The y axis measures proportional change in forest cover relative to the amount of forest cover prior to the intervention (with this baseline forest cover denoted as C0). The black lines trace out forest cover trends in the program (“treated”) areas, and gray shaded areas show the 95% confidence interval for the counterfactual forest cover trends implied by the effect estimates shown in Table 5 and Figure 4. The first five graphs show effects on forest cover attributable to deforestation (in which case, trajectories can never go above zero), the next five on forest attributable to either deforestation or forest growth. Effect estimates from Scullion et al. (2011) were not accompanied by standard errors and so we simply trace out the point estimate for the counterfactual forest cover trend.....	55

1. Background

1.1. Description of the problem

Natural forest preservation in the tropics, and thus in developing countries, is considered a key component of any effective effort to manage climate change. Forests serve as natural carbon sinks, which help to mitigate the effect of other carbon emissions. However, forest cover is being reduced and it is estimated that deforestation is responsible for 10-17 per cent of global carbon emissions. Gullison et al. (2007: 985) estimated that '[r]educing deforestation rates by 50 per cent by 2050 and then maintaining them at this level until 2100' would have a net impact of reducing global carbon emissions by an amount 'equivalent to nearly 6 years of recent annual fossil fuel emissions'. Similarly, in an essay that helped to inspire the REDD initiative, Santilli et al. (2005: 267) noted the importance of tropical natural forest protection for climate change, indicating that 'current annual rates of tropical deforestation from Brazil and Indonesia alone would equal four-fifths of the emissions reductions gained by implementing the Kyoto Protocol in its first commitment period [i.e., 2008-2012]'. Furthermore, Gullison et al. estimated that per unit costs of emissions reduction through forest protection, particularly if undertaken in tropical areas, may be substantially less than other approaches to emissions reduction, although these cost projections are based on assumption-driven, ex ante analyses rather than on ex post evaluations of tropical conservation programs.

In the context of global efforts to manage climate change, governments have coordinated forest conservation efforts under the Reducing Emissions through Deforestation and Forest Degradation (REDD) initiative since 2007. The REDD initiative consists of inter-governmental framework agreements to facilitate the protection of natural forests, with particular emphasis on natural forests in the tropics and, therefore, developing countries (that is, countries classified by the World Bank as low- or middle-income). The goal of REDD (now REDD+) is both to reduce carbon emissions resulting directly from deforestation and to preserve natural forests as carbon sinks so as to mitigate the effect of other carbon emissions on climate change (Intergovernmental Panel on Climate Change, 2007; Harris et al. 2012).

This has led to the implementation of various programs designed to reduce the amount of forested land converted to other purposes. Payment for environmental services (PES) programs is one type of intervention commonly implemented under the REDD+ umbrella and are widely implemented around the world as part of government strategies to manage forest loss and climate change. PES programs allow for direct exchange between those demanding 'environmental services' such as protection or rehabilitation of natural forests and those in a position to provide them locally (Forest Trends, Katoomba Group and UNEP, 2008; Wunder, 2005). Governments have applied PES strategies domestically for decades to manage forests and prevent irredeemable loss of valuable endemic forest resources. PES exist alongside 'decentralized forest management', 'community-based forest management' and 'protected areas' (that is, parks and reserves) as core components of government and privately led forest management efforts around the world (Angelsen, 2009).

Fundamental issues in policy debates over conservation strategies in developing countries include the extent to which conservation and poverty reduction goals conflict, how different conservation strategies fare in terms of such trade-offs, and the scope for 'win-win'

strategies that generate both significant conservation and poverty reduction benefits (Muradian et al., 2013; Sunderlin et al., 2005; Wunder, 2001, 2013). This review is organized around these issues. Two core questions arise. First, how might these potential benefits from natural forest conservation in the tropics be realized? Second, how do different approaches to natural forest conservation relate to the pursuit of poverty reduction goals? While tropical forests are appealing as targets for conservation because of their high carbon storage density and lower (in absolute terms) opportunity costs of conservation, they are located primarily in areas of low- and middle-income countries where poverty is a central concern (Deveny et al., 2009; Kremen et al., 2000; Sunderlin et al., 2005; Van Kooten and Sohngen, 2007). It is therefore crucial to understand whether conservation strategies require trading off on poverty reduction goals or whether there are strategies that allow for synergy in the pursuit of conservation and poverty reduction goals jointly. In this review, we address these questions with respect to PES programs.

1.2. Description of the intervention

1.2.1. Description of PES

At the most basic level, PES (also called ‘payment for ecological services’ or ‘reward for environmental/ecological services’) refers to voluntary accession to a contract to provide a well-defined environmental service (for example, maintenance of natural forest density in a designated area) in exchange for payment or other reward from a buyer entity. Our conception of PES encompasses what is sometimes referred to as ‘rewards’ for environmental services or ‘compensation’ for environmental services. Whether PES should be defined to include additional provisions is subject to some debate. Wunder (2005), in a sophisticated treatment of this definitional issue, defines an ‘ideal’ PES program as one that involves (i) such a voluntary exchange as well as (ii) the payment being issued conditional on delivery of the environmental service (as opposed, say, to being issued prior to and in expectation of delivery of the service) and (iii) the buyer entity being the immediate users of the environmental services.

Such an ideal form of PES is appealing in theory, as it would seem to define an ‘incentive-compatible,’ and therefore sustainable, approach to environmental protection. But we feel that this ideal form is ill suited to a review of PES programs as they have been implemented around the world thus far. As Wunder (2005; 2007) and Wunder, Engel, and Pagiola (2008) note, such ideal type programming is extremely rare in developing countries¹.

¹ Conventional economic theory provides possible explanations for this. To the extent that any local demand for environmental services exists (and it is not clear that it always will), the benefits from such services have the quality of public goods (Samuelson, 1954). Therefore, any single beneficiary would have an incentive to free ride on others’ purchase of environmental services, introducing the potential for market failure in the absence of government or other third party coordination (Salzman, 2010). Furthermore, if demand originates predominately among non-local or foreign entities, transaction costs may make direct contracting with local services providers impractical, again undermining the potential for market formation. Thus, government or NGO intervention is likely to be required to overcome market failure risks and organize the purchase of environmental services in developing country contexts. Governments or NGOs may find conditional payment to be sub-optimal in satisfying their manifold goals. For example, an NGO may have a goal of building trust with forest communities, in which case the NGO may find more appealing a “gift exchange” approach (Akerlof, 1982), where at least partial payment or administration of benefits is issued prior to service delivery, with the expectation that recipients will reciprocate by providing environmental services.

Taking this into account and in a manner consistent with the review by Wunder, Engel, and Pagiola (2008), we define PES as ‘voluntary accession to a contract to supply a well-defined environmental service in exchange for payment from a buyer entity,’ where payments need not be monetary but may come in the form of other material benefits and the ‘environmental service’ must involve the maintenance or rehabilitation of natural forests. ‘PES programs’ are actions undertaken by corporate or government entities to facilitate PES by establishing necessary legal frameworks (for example, by demarcating property rights) or connecting potential ‘buyers’ to potential providers of environmental services. While PES refers to a type of exchange that may be realized at any time, PES programs have clear start dates that allow in principle for an evaluation of their impacts.

By this definition, our review of the literature has found PES programs underway in many countries with large forest areas, including Bolivia, Brazil, Cameroon, China, Costa Rica, Honduras, Indonesia, Madagascar, Malawi, Mexico, Mozambique, Nicaragua, and Vietnam. PES program ‘inputs’ include funds to be used for the payments (e.g., from taxes, donor grants, or purchases of carbon offsets), staff that recruit service providers and manage contracts, and the potential service providers themselves and their land. The ‘outputs’ include contracts with service providers, hectares of land covered by such contracts, and payments issued in return for services. These inputs and outputs are intended to generate beneficial environmental and, potentially, welfare outcomes.

1.3. How the intervention might work

1.3.1. Main hypotheses

Figure 1 is a schematic representation of the theory of change that we evaluate with this review. We embed the causal relationships between PES and poverty/deforestation in Ostrom (2007)’s generic analytic framework for conservation dynamics. The framework defines the context in terms of the resource system, governance system, resource units, and resource users. In this review, the governance systems and resource users are the key areas of contextual variation that may moderate PES effects. Resource system (forest systems in developing countries) and resource units (forested land) are assumed to be of secondary concern once we condition on governance systems and resource users, with the latter understood as being potential agents of deforestation². The causal arrows in the diagram do not characterize all conceivable causal relationships, just the ones that we seek to test. We have drawn a causal arrow that flows from poverty to deforestation, and not the other way around. This does not mean that we assume no effects of deforestation on poverty. It is meant to clarify the particular mediating relationship that interests us in this review.

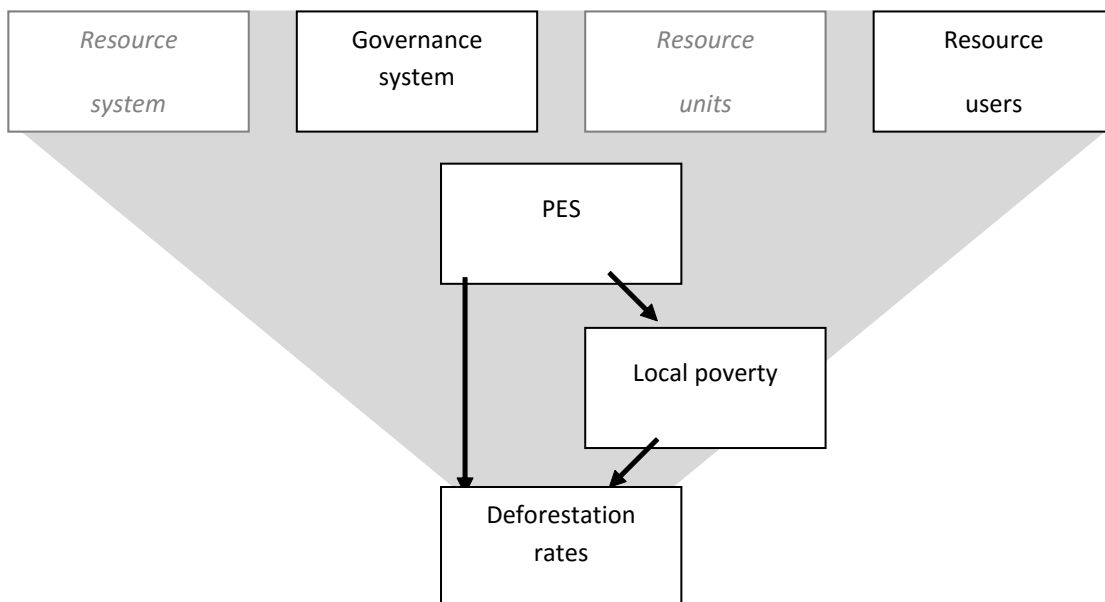
² This is mostly a semantic point: clearly aspects of the resource system will influence both deforestation rates and poverty conditions and will also affect the likely impact of PES programs. For example, whether there are highly valuable timbers or whether there are mining opportunities. However, resource systems factors such as these operate through resource users’ opportunity costs. Our argument is that conditional on opportunity costs, timber values and mining opportunities per se are of secondary importance. On the basis of economic theory, opportunity costs provide a sufficient statistic for incorporating such resource system factors into the analysis.

A crucial question for conservation programs in developing countries is whether there might be synergies between the pursuit of conservation goals and poverty reduction goals or not. Pagiola et al. (2005) argues that coupling poverty goals with environmental protection goals in conservation programming may be inefficient for reaching either type of goal, and that in many instances the two objectives are orthogonal to each other, if not in conflict. While poorer members of forest edge communities stand to gain the most from poverty alleviation programming, they may not constitute the greatest deforestation threat. Such individuals may have relatively little means or incentive to engage in deforestation relative to large-scale farmers or logging interests. If so, making poverty alleviation in forest edge communities a priority may imply inefficient targeting of resources if the goal is the biggest conservation payoff (Wunder, 2005: 12-14).

There is a moral reason to couple poverty relief with conservation, although it does not presume the possibility of synergy: such a coupling would be imperative if conservation disrupts livelihoods of forest community members by limiting their ability to exploit resources for productive purposes, whether by themselves or as hired labor (Agrawal and Benson, 2011; Angelsen and Wunder, 2003; Chomitz, 2007: Ch. 3; Edwards et al., 2011; Porras, 2010; Wunder, 2005).

Arguments for synergy include those based on a sustainable livelihoods and political logic. With respect to sustainable livelihoods, the classic study by Vandermeer and Perfecto (1995) detailed tropical deforestation threats arising from forest edge communities' abandonment

Figure 1: Illustrating the theory of change



of sustainable forest use practices in the face of various pressures from commercial agriculture. Poverty relief for such communities is proposed as a way to arrest such dynamics. Politically, conservation strategies may be made more viable and effective if coupled with poverty alleviation. If PES programs target only the interests of large-scale commercial enterprises, the result may be to exacerbate local inequality. Moreover, attaching poverty alleviation goals to conservation programs may help to minimize risks of hostilities, local level subversion, and corruption (Mapedza, 2006).

Based on this theoretical discussion, the two most basic hypotheses that we wish to test are as follows:

- H1: PES reduces deforestation rates
- H2: PES has non-negative impacts on local poverty levels.

The focus on non-negative, as opposed to “positive” impacts per se, reflects a primary concern to ensure that policy interventions do no incidental harm in association with ultimate goals, which in this case are taken to be reductions in deforestation.

Beyond these basic effects, we are interested in the possible mediating role of poverty conditions for deforestation outcomes. The sceptical take is that the two dimensions are orthogonal or even conflicting. The ‘synergy’ position is that poverty consequences of conservation policies mediate effects on deforestation. For PES, the nature of the dilemma that pits attending to distributional concerns against targeting major agents of deforestation likely depends on two factors. First are levels of local inequality in terms of holdings and vulnerability due to cessation of deforesting activities (part of the ‘resource users’ context). Second is the political position of those who stand to lose from cessation of deforesting activities relative to those who stand to gain from PES. Thus, we have the following hypotheses:

- H3: The more a PES program functions to relieve poverty, the stronger its impact will be on reducing deforestation.
- H4: PES deforestation reduction impact is negatively moderated by prevailing levels of local inequality in holdings of forested property and vulnerability due to cessation of deforesting activities.

1.3.2. Unintended consequences and moderating factors for PES

The four hypotheses stated so far reflect the main policy interests motivating this review, but they do they reflect consensus opinion on how PES may work. Various moderating factors and potentially unintended outcomes need to be considered. The justification for PES, ostensibly, is that without intervention, benefits of forest protection are external to those who would contribute to deforestation. PES programs thus harness and redirect the value of such externalities in the form of payments to those who would otherwise contribute to deforestation (Wunder, 2005; Angelsen 2010).

In principle, a PES arrangement operates as a standard performance-based contract, whereby upon performance, in terms of forest protection or rehabilitation, payments are issued (Ferraro, 2011). A number of conditions are necessary for such an arrangement to work. Payments must be targeted toward those whose activities significantly affect

deforestation rates, the payments must be sufficient to overcome opportunity costs of conservation, and the 'sellers' must be induced to carry out the conservation service rather than pocketing the payment and continuing with deforestation (Ferraro, 2011).

These conditions may fail if institutional conditions or even cultural conditions (e.g., 'a payment culture,' Muradian et al. 2013; Wunder, 2013) are not right. Those designing the program may have inadequate knowledge or capacity to target properly or to set appropriate payment levels³. Constraints on PES buyers' ability to monitor and sanction may allow would-be sellers of conservation services to get away with pocketing benefits without actually reducing deforestation. These institutional conditions are captured by the 'governance system.' A PES buyer's ability to monitor and sanction will depend on local public administration, law enforcement capacity, as well as levels of corruption.

Poverty alleviation effects of PES will depend on the targeting of the program and whether local institutions represent the interest of the poor and therefore provide accountability mechanisms to ensure that benefits accrue to the poor (Corbera et al., 2009). Average incomes may rise, for example, but these gains may be concentrated among the non-poor, in which case poverty levels may be unchanged.

It is also possible for PES to have perverse or unintended negative effects. For instance, cessation of deforestation may reduce demand for labor from poor households or otherwise infringe on livelihoods of the poor, leading to welfare losses. PES may increase the value of land and result in more powerful groups displacing poorer households so as to gain control of the land that the poor occupy, again resulting in welfare losses among the poor (Landell-Mills and Porras 2002; Langholz et al., 2000). Finally, the 'commodification' of forest could erode people's sense of the intrinsic value of forests; this could make custodians of forested land more receptive to bids proposing commercial conversion of forests that are likely to be more lucrative than conservation contracts (Muradian et al., 2013, though see also Wunder, 2013).

1.3.3. Moderator hypotheses

We can state these points from the two sections above in terms of moderator hypotheses. The hypotheses are as follows:

- H5: PES deforestation reduction impacts are positively moderated by the level of local administrative and enforcement capacity.
- H6: PES deforestation reduction impacts are negatively moderated by levels of corruption in government.

In testing these hypotheses, we control for variations in the design features of PES programs. PES programs vary by the size and terms of the payments offered. Differences

³ Indeed, as Boerner and Wunder (2008) and Grieg-Gran (2008) demonstrate, ex ante valuation of opportunity costs of forest conservation is not a trivial undertaking. For example, in valuing the opportunity costs in two sites in Brazil, Boerner and Wunder combine production and price data with forest loss projections to derive valuations that could be used to scale payments. Along similar lines, Gregersen et al. (2010) further problematize opportunity cost analyses by pointing out that some opportunity costs may be illegal or based on informal markets (e.g., illegal timber trade or slash and burn farming), making it especially difficult to establish clear opportunity cost benchmarks.

across programs will reflect policy-makers' adaptation to contextual factors. In our analysis, we study how contextual variables, and in particular the moderating factors discussed above, affect the PES design.

1.4. Why it was important to do this review

While the environmental science is clear in characterizing the potential gains from forest conservation (Santilli et al., 2005; Gullison et al., 2007), it remains for social scientists to provide insights into how institutions and incentives may be arranged to realize such potential (Gibson et al., 2000). Realized impacts may depart substantially from hypothetical projections, in which case the latter on their own are not a reliable guide for policy. Evidence from case studies of PES programs is inconclusive about the effectiveness of such programs for forest protection; this may reflect how the implicit theories used to design PES programs have failed to account for local structural and institutional context (Angelsen, 2010; Ferraro et al., 2012; Gibson et al., 2000; Tacconi, 2007; Wunder, 2005).

Typically evaluation approaches in this field estimate the worthiness of conservation programs on the basis of elicited valuations of environmental services combined with hypothetical projections of the services that a program is supposed to deliver. As Ferraro et al. (2011) and Ferraro (2011) argue quite convincingly, there is a need to move toward credible estimation of the effects of conservation programs. Efforts to apply counterfactual analysis to assess the effects of environmental programs have been rather limited to date, but studies using quasi-experimental approaches do exist. These are currently scattered in the academic and grey literature, with no comprehensive synthesis available to date.

This review complements a number of other systematic reviews assessing the evidence on interventions considered under the REDD+ initiative and other efforts to reduce deforestation. Bowler et al. (2010) assess the effects of community based forest management on environmental and human welfare outcomes. Geldman et al. (2013) assess the effects of protected areas on environmental outcomes, while Pullin et al. (2013) focus on the human welfare outcomes of protected areas. Finally, Samii et al. (forthcoming) was conducted in parallel to the current project and focus on the effects of decentralized forest management on environmental and human welfare outcomes.

There is a range of recent, related review studies that have had similar goals as this review, and so it is important to clarify our added contribution. Pattanayak et al. (2010) review theoretical motivations for forest-oriented PES and findings from eight quasi-experimental studies and 18 case studies, but do not apply the replicable search and synthesis methods of a systematic review. Wunder et al. (2008) review evidence on distributive effects of forest-oriented PES programs from case studies, but do not provide quantitative synthesis. The volume edited by Angelsen (2009) contains chapters that describe varieties of forest conservation policies, including forest-oriented PES programs, but these reviews do not adopt the replicable search and synthesis methods of a systematic review.

2. Objectives

The overall objective of this review is to assess the evidence on the conservation and poverty impact of PES programs and to assess the extent to which the poverty impact of such programs in turn affects the extent to which conservation benefits are realized. Doing so is important for moving the debate outlined in section 1.1 beyond theoretical discussions and into better-informed, evidence-based discussion (assuming relevant evidence can be found). More specifically, we seek to test the hypotheses set forth above, with hypotheses H1 and H2 being of primary interest. Hypotheses H4 through H6 are of secondary interest and in testing them we seek to evaluate how institutional and social conditions (namely, inequality, institutional capacity, corruption, and democratic accountability) moderate the impact of PES programs. Our strategy for selecting studies will be targeted toward testing the four primary hypotheses as rigorously as possible. Table 1 relates each hypothesis to the types of evidence we will need. Such an assessment of impacts does not necessarily provide the basis for a full cost-benefit analysis of PES programs. We acknowledge this limitation and propose that follow-up work should focus on filling in the cost side of the equation as a complement to the analysis that we provide in this report.

Table 1: Questions and types of evidence needed for the review

Hypothesis	Type of evidence
Main Hypotheses <ul style="list-style-type: none"> H1: PES reduce deforestation rates. H2: PES have non-negative impact on local poverty levels. 	Quantitative data on forest conservation and host community poverty outcomes for sites with PES and sites that constitute a plausible counterfactual. Qualitative accounts of whether the interventions operated as planned.
Mediator Hypothesis H3: The more a PES program functions to relieve poverty, the stronger will be its impact on reducing deforestation.	Quantitative estimates of both poverty and deforestation impacts from PES for at least a subset of cases to assess co-variation between the two types of impact. Qualitative accounts of whether poverty benefits (disruption) contributed to compliance (non-compliance) and effective (ineffective) functioning of PES programs.
Moderator Hypotheses <ul style="list-style-type: none"> H4: PES deforestation reduction impact is negatively moderated by the level of local inequality. H5: PES deforestation reduction impacts are positively moderated by the level of local administrative and enforcement capacity. H6: PES deforestation reduction impacts are negatively moderated by the level of corruption in government. 	Quantitative measures of local inequality, local capacity, corruption, local democratic accountability, and opportunity costs of conservation borne by forest communities for each study to assess covariation between these measures on the one hand and deforestation and poverty on the other. Qualitative accounts of how issues related to inequality, local capacity, corruption, or local democratic accountability affected the functioning and effectiveness of given PES programs.

3. Selection Criteria

Our selection criteria are summarized in Table 2. Details are given in the following subsections.

3.1. Participants

This review includes only studies that focus on either (i) deforestation outcomes in forest areas in developing countries or (ii) poverty conditions of forest-dwellers and populations residing in communities that are proximate to natural growth forest areas in developing countries. ‘Forest’ is defined as per the United Nations Food and Agricultural Organization Global Forest Resources Assessment:

Land spanning more than 0.5 hectares with trees higher than 5 meters and a canopy cover of more than 10 percent, or trees able to reach these thresholds in situ. It does not include land that is predominantly under agricultural or urban land use. (Food and Agricultural Organization, 2010: 6)

‘Developing countries’ are those classified as lower income, lower middle income, or upper middle income by the World Bank in the year of the initiation of the program under study.

3.2. Interventions

The review includes studies of PES programs. The requirements for a program to be considered a PES program are that there is a clear start date when either payments or rewards are themselves offered to individual or corporate property holders to maintain or rehabilitate (for example, via planting endemic species) natural forests, or institutions are established to facilitate such offers. We allow for those offering the rewards (the ‘buyer entity’) to be either public or private actors, and we allow for payments to be made in a manner that is either conditional or in a manner that is in advance (and therefore not necessarily conditional) on the fulfilment of the prescribed maintenance or rehabilitation. These differences are noted in our characterization of the design of each PES program below.

3.3. Outcomes

Outcomes of interest are (i) deforestation or (ii) poverty conditions of forest-dwelling communities. Similar to what Bowler et al. (2010) discovered, in our selected studies researchers varied in the precise metric that they used for deforestation impacts, including differences in operational definitions for deforestation or degradation and different types of data sources---for example, on-the-ground point samples or remote sensing samples from satellite or fly-over imagery (West, 2009; Achard and Hansen et al., 2013). We accepted whatever measure was used for the outcomes of interest as presented by the authors.

We sought to assess poverty impacts in terms of impacts on consumption, income, or income potential for members of forest communities residing below or just above the consumption-based, two-dollar per day purchasing power parity absolute poverty line (Ravallion et al., 1991). Such outcomes are typically assessed using household economic surveys or administrative data on consumption, food security, employment, or access to productive assets (Deaton, 1997). In the absence of such fine-grained data, we sought to look at studies that measure differential consumption or income impacts for ‘poor’ versus

'non-poor' households or communities. Again, we accepted whatever measure was used for the outcomes of interest as presented by the authors.

We also sought to pay attention to the potential impact of in- or out-migration on poverty outcomes. If a program causes outmigration among the most poor, then the resulting poverty level in the area may be less than was the case before the program. However, it would be inappropriate to take this to mean that the program helped to alleviate poverty.

Finally, we were particularly interested in identifying unintended effects of forest conservation programs on local poverty conditions. We also took note of whether studies accounted for spill-over effects such as deforestation 'leakage' or 'slippage' (Wu et al., 2001). Failure to account for such spill-over may result in a biased interpretation of the impact of a program.

3.4. Study Types

Table 1 above sketched out the types of quantitative data and qualitative evidence we included in this review. We prioritized identifying rigorous studies that address hypotheses H1 and H2. For quantitative synthesis, we sought well-designed experimental or quasi-experimental studies that use robust methods to construct approximations to the counterfactual for the areas or individuals subject to a PES program. We then made comparisons between outcomes in the 'treatment' group and outcomes in the approximation to the counterfactual for the treatment group.

We accepted for quantitative synthesis only (a) randomized studies or (b) quasi-experimental studies that employ strategies for causal identification with clearly delineated treated and comparison areas and use some method for removing biases due to non-random assignment of treatment. Such methods include: regression adjustment, difference-in-differences estimation, instrumental variables regression, fixed effects regression, regression discontinuity, matching, or inverse-propensity-weighted estimation. While application of such a method is sufficient for inclusion in our study, we appreciate that not all studies apply methods for causal identification with equal rigor and therefore we assessed the quality of all included studies (below we discuss the tools we used to assess study quality).

Quantitative studies that were excluded were those that failed to establish a credible approximation to the treatment group counterfactual. This included studies that relied exclusively on uncontrolled before-after comparisons or failed to adopt any of the above-mentioned methods of analysis to correct for selection bias and confounding. Qualitative data are used in the synthesis to provide descriptions and context for interventions that are included in the quantitative synthesis. Such data were drawn from the quantitative studies themselves as well as qualitative studies that cover the same programs or settings as the quantitative studies.

Table 2: PICOS inclusion criteria

Type	Criteria
Participants	Forest areas or forest communities in developing countries.
Interventions	PES programs.
Comparisons	“PES versus no PES contemporaneous counterfactual”
Outcomes	Deforestation or poverty among forest communities.
Study types	Quantitative studies providing a robust counterfactual via randomized experiment or quasi-experiment or qualitative study with clear research objectives, original analysis, explanation of methods, and seeking to contribute to the academic conservation or social science literature.

4. Search Strategy

4.1. Electronic searches

Our search strategy was developed after initial scoping exercises with a Campbell Collaboration information retrieval specialist. We searched the set of databases, specialist websites, and search engines that Bowler et al. (2010: 55-56) searched as well as others identified to possibly contain relevant content⁴. Our list of sources searched is given in appendix 11.1 below.

Our search strings included the following key words:

("pay*" OR "reward*" OR "incentiv*" OR "compensat*") AND ("forest" OR "deforest*" OR "ecol*" OR "ecos*" OR "environment*" OR "conservation")

To these keywords we also applied a lower- or middle-income filter based on the Cochrane EPOC filters (<http://epocoslo.cochrane.org/lmic-filters>). The search strategy was adapted for individual databases. An example of a full search strategy is included in appendix 11.1.4 below.

Some of the databases considered (for example, IDEAS, RUPES, and JSTOR) included search results for non-English language studies even when using English search terms and keywords. The relevance of such search results was reviewed by native language speakers (the authors were able to cover French, Spanish, German, and Bahasa Indonesia). Ultimately, only English language studies met our inclusion criteria. We did not impose any date restrictions.

4.2. Other Searches

We carried out hand searches of (i) key journals in relevant fields as listed in the appendix, using publisher search engines and (ii) references cited in papers accepted for review as well as in review papers or thematically relevant papers identified during the search. We had members of our advisory group and the specialist agencies listed in the appendix below review our search results to ensure that important studies were not missing from our search results.

⁴ We apply the same strategy of reviewing only the first 100 hits for internet search engine results (but not academic database results), given that search engines typically return many thousands of results.

5. Data Collection and Analysis

5.1. Data collection and analysis

5.1.1. Selection of studies

The review team applied the PICOS inclusion criteria listed in Table 2 in three stages: first to titles to remove spurious citations, then to abstracts, and finally to full texts. For all stages, we maintained an account of the number of studies excluded, and the reasons for exclusion, by tracking references in an Endnote database. In the full text stage, excluded studies were tagged in terms of the PICOS criteria that were violated. All screening was done by two independent reviewers from the research team, with disagreements resolved by a third reviewer from the team. To ensure consistency in selection procedures, multiple reviewers reviewed a sample (of 50, for example) of citations and consistency was assessed. If agreement rates were below 90 per cent, we addressed any inconsistencies in interpretation of the criteria to assure at least 90 per cent rates of agreement.

5.1.2. Data extraction and management

For studies eligible for inclusion, we collected data on the study characteristics, findings, and moderators using a coding form (see appendix section 9.6). The data were double entered into Microsoft Excel by the review team. While it would be ideal to have data on moderator variables measured at the level of the regions in which the programs under study are applied, such data were not typically available. Therefore, we obtained data on the moderator variables using the relevant country-level indicators from the World Bank Governance Indicators. In the end, because of the low number of countries represented, there was little that we could do with these moderator variables.

5.1.3. Assessment of risk of bias in included quantitative studies

Risk of bias was assessed based on the guidance of the IDCG Risk of Bias Tool (version March 2012)⁵. We appraised studies according to the following criteria:

- Avoiding selection bias due to non-random assignment, non-exogenous source of quasi-experimental variation in assignment, no adjustment for differences in baseline measurements: We assessed this on the basis of whether or not the study worked with a source of exogenous treatment assignment.
- Avoiding confounding bias due to lack of control for key confounders: Based on an initial reading of the studies, we concluded that key confounders included variables related to land quality, socio-economic conditions (namely, livelihoods, living standards, and access and size of markets for agricultural producers), and accessibility of treated land areas. We assessed whether studies accounted for all three types of confounders.
- Avoiding motivation bias from measurement strategies that may be tainted by subjects' interest in presenting themselves in a positive light or telling researchers

⁵ We improved our risk of bias assessment over what we had proposed in the protocol to account for more specific nuances of the studies under consideration.

‘what they want to hear’: This was assessed as being satisfied if study conclusions were drawn from effects estimated on non-self-reported data or data based using other measurement strategies that reduce motivation biases.

- Accounting for potential bias due to spill-overs: We assessed whether studies either evaluated units that were insulated from spill-over or, in case where spill over was a likely concern, tried to estimate the extent to which spill-over may bias naïve comparisons.
- Free of selective outcome reporting and analysis fishing: We assessed whether studies clearly omitted results that might undermine the conclusions of the study or drew conclusions on the basis of methods that showed high potential for specification search.
- Appropriate statistical inference due to proper calculation of standard errors and confidence intervals.

We coded each study on the basis of whether they clearly satisfied each of these conditions (coded as ‘yes’), clearly failed to do so (‘no’), or whether it was impossible to judge (‘unclear’).

5.1.4. Measures of treatment effect

For effects on forest cover, whenever possible we standardize effect estimates to annual forest cover change rates following the proposals of Puyravaud (2003). For effects on material welfare and poverty effects, we used percentage change over estimated average counterfactual outcome (e.g., for income effects, percentage change in income relative to the average income of the control group). Section 9.3 of the appendix provides the precise calculations for these standardized measures and associated standard error approximations.

When multiple estimates were presented in a given study, we first tried to select the estimate that posed the lowest risk of bias. For studies relying on ‘conditional independence assumptions’ and using multiple regression or matching, this would be the estimate that either controlled-for or achieved the best balance on the largest set of pre-treatment covariates⁶. When there was no clearly defensible way to identify the single estimate in a study with the least risk of bias, we extracted all estimates and then performed our synthesis with the mean of the different estimates as well as the mean of the standard error estimates. This approach does not account for the dependence of the different effect estimates, although it avoids pitfalls in the use of standard approaches that assume independence⁷.

6 As discussed in Lawry et al. (2013), in some cases, adding more covariates can actually increase the bias of an estimate, but this is something that is impossible to judge from the data.

7 Initially we had use an inverse variance weighted averaging approach for synthesizing the different effect estimates. But, as a reviewer astutely pointed out, such an approach ignores the dependence between measures and results in synthesized standard errors that become artificially small as one increases the number of estimates. Our approach to using the mean of the effect estimates and standard errors was proposed as the least misleading way to synthesize effect estimates when there is no clear way to select one minimally biased estimate.

Some of the studies that we included examine the same program, however the estimates that they present cover different time periods, cover different regions, and use independent data sources. As such we treat these as distinct (and statistically independent) estimates.

5.1.5. Unit of analysis issues

When the unit of analysis was at a lower level of aggregation than assignment units, standard error calculations should account for the attendant ‘clustering’. We checked to be sure that this was done. In cases where it was not, we noted it in our risk of bias assessment and while we sought to correct them using standard formula in cases where the relevant problems arose the information was not available to do so.

5.1.6. Dealing with missing data and incomplete data

When studies did not report on endpoint or intermediate outcomes, the study authors were contacted to determine whether such outcome data did in fact exist and whether estimates could be produced. However, we did not receive data from any authors that would allow for the construction of effect estimates that went beyond what appeared in the original studies.

5.2. Data synthesis

5.1.7. Quantitative Synthesis

Our plan for a quantitative synthesis was guided by the hypotheses listed in Table 1. The ‘main hypotheses’ (H1 and H2) require a synthesis of basic effect estimates on deforestation and welfare or poverty. For each hypothesis, the following conditions had to apply for a statistical meta-analysis to be justified (adapted from Wilson et al., 2011): i) more than two studies meeting the quantitative inclusion criteria with effect sizes for common outcome constructs AND ii) effects measured against similar comparators.

Only for the effects of PES on deforestation were these conditions met. We thus computed an average overall effect as a weighted average that accounts for the imprecision of each effect estimate. The estimate was constructed using a random effects model fit via empirical Bayes in the metaphor package for R (Viechtbauer, 2010). In our forest plot for PES deforestation effects we display the synthetic random effects means along with their 95 percent confidence intervals (displayed as a black diamond).

The limits of the evidence base prevented further meta-analyses. In our protocol, we proposed a meta-regression approach for testing the moderator and mediator hypotheses. We could not implement this approach for lack of studies. Rather, we were forced to rely on qualitative information relevant for the included studies to comment on, rather than test, the moderator and mediator hypotheses. For similar reasons, we could not implement quantitative analyses of publication biases.

For the most part, our quantitative synthesis is limited to tables of effect estimates and narrative discussions of trends in the size and direction of the effects reported by the studies. The narrative discussion highlights issues related to modes of measurement, nature of comparators, as well as moderator conditions that should be taken into account when comparing the different effect estimates. We also provide a critical assessment of methods that have been employed and provide concrete recommendations for how rigorous and comparable evidence might be generated in future research.

6. *Use of qualitative data*

We extracted qualitative information from both the included quantitative studies as well as qualitative studies that covered the same types of programs and contexts (defined by our moderator variables) as our quantitative studies. We use such qualitative data to establish that conditions recorded in quantitative data are correctly interpreted and that hypothesized, but difficult to measure, chains of events do in fact occur in linking explanatory factors to outcomes (Collier, 2011; Vajja and White, 2008). Our strategy was to search on hypothesis-specific keyword word stems in the articles for the mediator and moderator hypotheses outlined above. We used these search results to localize content that may be relevant to our hypotheses. We extracted qualitative accounts or conclusions that were relevant to each of the hypotheses, and used these to provide insights in our narrative discussion. The keyword word stems that we used included the following:

- H3a & H3b: poverty or welfare.
- H4: equal, fair, rights, or property.
- H5a & H5b: capacity, monitor, technical, difficult, or governance.
- H6a & H6b: corrupt, rent, elite, capture, or profit.
- H7: politic, voice, democrat, participat, mobiliz, or accountab.

6. Results

Our search for qualifying studies followed the process presented in Figure 2. This search process identified 1382 articles on PES using the search terms described in Section 4.1. Screening of abstracts had us narrow this to 149 PES studies to be screened at full text. Screening of full text papers reduced this first to a set of 11 quantitative studies and one qualitative PES study that met our inclusion criteria. We then conducted a second targeted search for other relevant and methodologically adequate qualitative studies that our initial search did not recover. We did this second targeted search by identifying any qualitative studies referenced in the bibliographies of the accepted quantitative studies, checking the websites of the quantitative study authors to see if they had produced complementary qualitative research, and then searching in the same databases as in the initial search, using as search terms the names of the programs that were being evaluated in the quantitative studies. This yielded eight new qualitative studies of PES programs assessed in the quantitative studies, or in studies from the same contexts as those studies. Therefore our final set was 11 quantitative and nine qualitative PES studies. Appendix section 9.7 provides information on studies that were excluded at the full text review stage. Tables 3 through 6 provide characteristics of the included studies, grouped by program.

Figure 2: Study search process

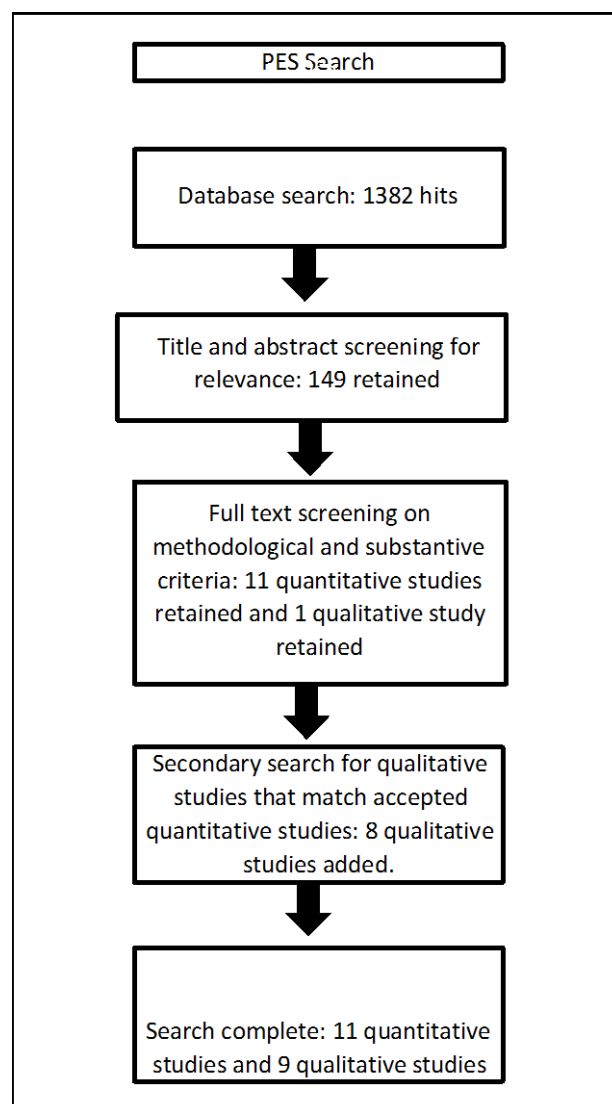


Table 3: Quantitative PES study characteristics (grouped by program)

Program	Studies	Treatment Group	Time Period	Counterfactual Approximation	Forest Conservation Outcome	Welfare Outcome	Method of attribution	Sample
China SCLP	Liu et al. 2012	2001-2004 SCLP households	2001-2004	regression adjusted households without SCLP		Natural log of total household income	multiple regression	Approximately 2000 rural households providing panel data over 1995-2004
Costa Rica PSA	Amiagada et al. 2012	1997-2005 PSA conservation contract farms in Sarapiquí	1997-2005	matched and difference-in-difference adjusted eligible farms in Sarapiquí without PSA	hectares of forest cover		matching and multiple regression	202 farms in Sarapiquí (including 50 PSA farms)
Costa Rica PSA	Amiagada et al. 2011	1998-2005 census tracts covered by PSA conservation contracts and on the common support with matched non-PSA tracts	1998-2005	nearest-neighbor matched census tracts without PSA (best covariate balance estimate) with multiple regression adjustment	hectares of forest deforested and hectares of forest cover		matching	8,073 census tracts (including 1,050 PSA tracts)
Costa Rica PSA	Robalino and Pfaff (2013)	1997-2000 1kmX1km forested PSA conservation contracts pixels	1997-2000	bias-adjusted matched forest pixels without PSA	percent of forest area deforested		matching and multiple regression	50,000 1kmX1km locations randomly sampled from across Costa Rica
Costa Rica PSA	Robalino et al. 2008	2000-2005 1kmX1km forested PSA conservation contracts pixels	2000-2005	bias-adjusted matched forest pixels without PSA	percent of forest area deforested		matching and multiple regression	50,000 1kmX1km locations randomly sampled from across Costa Rica
Costa Rica PSA	Sanchez-Azofeifa et al. 2007	1997-2000 grid cells covered by PSA conservation contract	1997-2000	regression adjusted grid cells without PSA	probability that area was deforested		multiple regression	2,021 5kmX5km grid cells from across Costa Rica
Costa Rica RISEMP	Garbach et al. 2012	2002-2008 small/medium cattle farm properties in Esparza covered by PSA and technical assistance	2002-2008	convenience sample of local non-PES farmers	number of silvopastoral practices adopted		multiple regression	101 farms (including 35 PSA farms)
Mexico PSAH	Aliz-García et al. 2012	2004-2006 PSAH properties	2004-2006	bias-adjusted matched properties rejected in 2004 or online for PSAH for 2006	percent of forest area deforested		matching and multiple regression	633 parcels of which 371 were PSAH recipients and the rest were either rejected in 2004 or online for 2006
Mexico PSAH	Scullion et al. 2011	2000-2003 PSAH or FIDECOAGUA properties	2000-2003	local properties not under PSAH or FIDECOAGUA	hectares of pine oak forest cover and cloud forest cover		difference-in-differences	GIS polygons in Coatepec region
Mexico MBCF	Honey-Roses et al. 2011	2000-2009 MBCF properties in Monarch Butterfly Biosphere Reserve area	2000-2009	spatial matched properties without MBCF	percent of forest area still covered by forest and percent of forest area undisturbed		matching and multiple regression	9,441 GIS polygons in Monarch Butterfly Biosphere Reserve
Mozambique NCCL	Hegde and Bull 2011	2003-2006 NCCL participant households in project region	2003-2006	nearest neighbor matched local non-participant households		Household consumption expenditure and cash income in MTS	matching	290 households (including 96 participating households)

Table 4: Qualitative PES study characteristics

Program	Studies	Study Type	Data used
Costa Rica PSA	Arriagada et al. 2009	Analysis of participation in the program	Informal interviews and farmer surveys
Costa Rica PSA	Cole 2010	Analysis of survey data	Informal interviews and farmer surveys
Costa Rica PSA	Pagiola 2008	Political economy analysis and macro assessment of impact	Official documents and administrative records
Costa Rica PSA	Pagiola et al. 2005	Analysis of participation in the program and hypothetical analysis of impacts	Official documents and administrative records
Costa Rica PSA	Zbinden and Lee 2005	Analysis of participation in the program	Administrative records and household surveys
Mexico PSAH	Honey-Roses et al. 2009	Process evaluation	Administrative surveys, remote sensed data, and field samples
Mexico PSAH	McAfee and Shapiro 2010	Process evaluation and political economy analysis	Official documents, administrative records, and informal observation of policy formulation processes
Mexico PSAH	Munoz-Pina et al. 2008	Process evaluation and political economy analysis	Official documents and administrative records
Costa Rica PSA & Mexico PSAH	Wunder et al. 2008	Qualitative evidence synthesis	Synthesis of results from case studies

6.1. characteristics of included studies

The evidence base for the effects of PES on deforestation and poverty is extremely thin and these studies have methodological shortcomings. We identified a handful of high quality studies, which cover a small number of programs and contexts. Few of these studies provide insights on the intersection of forest conservation and poverty, and the moderating effect of the social and institutional context. Table 7 below display various design features for the programs evaluated in the quantitative synthesis. Section 9.2 in the appendix provides more detail on each of the programs.

Table 3 describes the eleven studies that met our quantitative inclusion criteria. The nine studies that provide evidence on deforestation effects cover only four different programs in two countries: Costa Rica's Pagos por Servicios Ambientales (PSA) program⁸, Costa Rica's implementation of the Regional Integrated Silvopastoral Approaches to Ecosystem Management (RISEMP) program, Mexico's Pago por Servicios Ambientales Hidrológicos (PSAH) program⁹, and Mexico's Monarch Butterfly Conservation Fund (MBCF).

The two studies that provide evidence on poverty effects cover two countries and programs: China's Sloping Land Conversion Program (SLCP) and Mozambique's Nhambita Community Carbon Livelihoods (NCCL) program (see appendix section 9.2 for descriptions). We did not identify any studies that provide evidence on both deforestation and poverty effects for a common program, which prevents us from carrying out some of the quantitative analyses that we hoped to do on how poverty effects might in turn mediate deforestation effects.

⁹ The set of studies on PSA examine different time periods and regions in which the program was implemented, so they are not redundant.

6.1.1. Risk of bias in included studies

In addition to the small number of studies, the evidence base suffers from methodological shortcomings. Figure 3 shows the results of our risk of bias assessment, summarised for all included studies (study by study risk of bias assessment is available in appendix 11.8). We did not identify any experimental studies, and only one study made use of a source of a plausibly exogenous variation: Alix-Garcia et al. (2012) sampled 'matched control' parcels from properties that were idiosyncratically excluded from the first PSAH cohort but admitted to a subsequent cohort. This was the only study that ensured that the 'control' parcels included in the analysis could verifiably be assumed to have some chance of having been treated by the program based on the expressed interest of the parcel owners, thereby reducing concerns about self-selection bias. All other studies required that this assumption be taken on faith.

While all studies performed some kind of confounder control, many failed to include the full combination of forest land quality, socio-economic conditions, and conditions determining accessibility to markets that are often associated with both PES take-up and PES impact (the importance of all three factors for both take-up and forest cover trends were demonstrated across the studies themselves). It is reasonable to assume that PES programs tend to be applied systematically to parcels that landowners have no intention to deforest. If research designs fail to adequately control for such selection effects, estimates obtained on forest conservation effects will be biased upward.

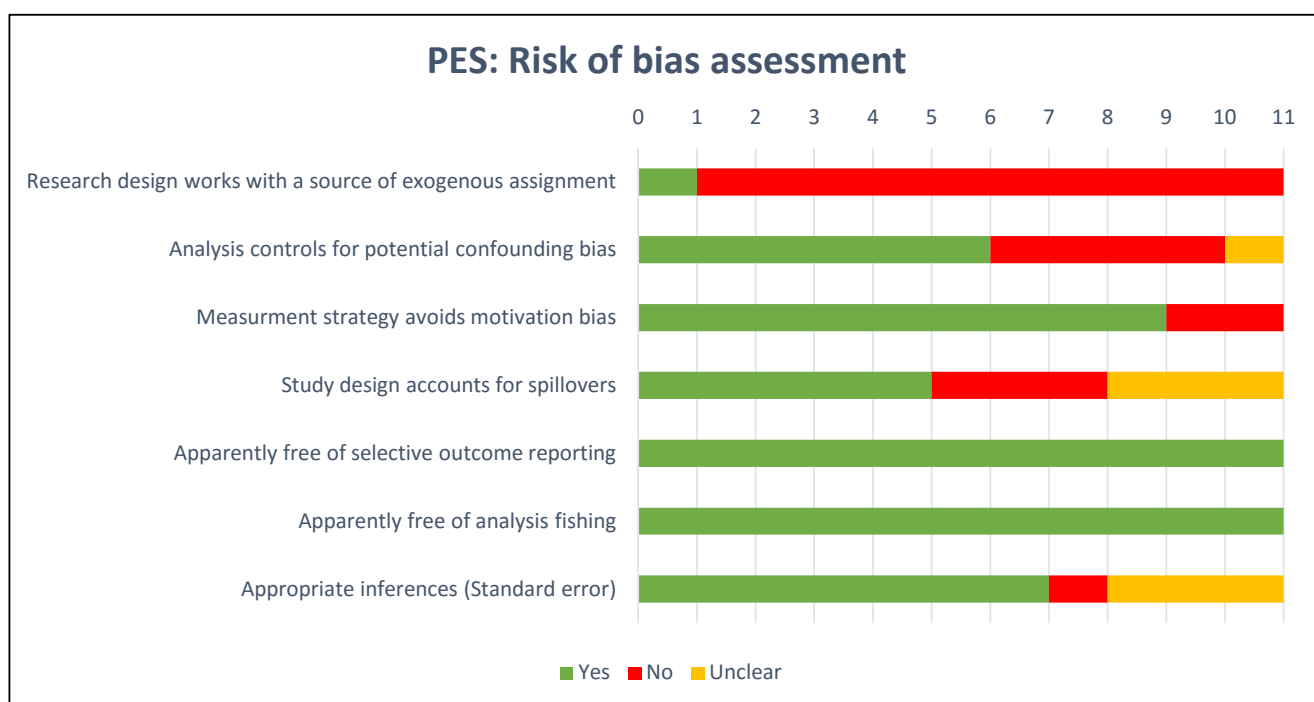
A majority also failed to give explicit attention to the issue of spill-over ('leakage' or 'slippage' when speaking of deforestation). Again, this will have the tendency of biasing upward estimates of program effects: deforestation displaced onto non-program parcels will be mistakenly interpreted as an approximation of what would have happened with no program. Therefore if opportunities seized by participants reduce opportunities available for non-participants, then their welfare may be worse than would be the case with no program, in which case the non-participants do not provide for a valid approximation to the no program counterfactual¹⁰.

Methods used by authors for statistical inference (standard errors, confidence intervals) were also problematic in some cases. For example, the study by Scullion et al. (2011) failed to provide any inferential statement on the effect estimates¹¹. In general, then, the methodological shortcomings of the evidence base are quite severe. For observational studies, the Alix-Garcia et al. (2012) provides a model that others ought to emulate. A move toward experimental studies would be helpful, and this is a point that we discuss in more detail in our conclusions below.

10 Alix-Garcia et al. [2012] find that local 'slippage' of this variety is probably minor, however: they estimate that naïve comparisons probably overstate PSAH program impacts by about 4 per cent.

11 The apparent rationale may be provided that the study used complete population data and therefore there was no uncertainty, sampling or otherwise. Such arguments misunderstand the nature of causal effect estimation: what is 'missing' in a causal analysis is data on counterfactual outcomes. Proxies from 'control' areas serve to fill in this missingness for the purposes of an analysis that estimates the 'treatment on the treated,' but the extent to which such a proxy departs from the true counterfactual is a source of uncertainty that needs to be accounted for (Abadie et al. 2014).

Figure 3: Risk of bias assessment



6.2. characteristics of included studies

Keeping in mind the weakness of the evidence base, Table 5 and Figure 4 display the estimated effects on forest cover outcomes for PES programs from the quantitative studies that qualified for inclusion. We converted estimates of forest cover effects to effects on the annual forest cover change rate (following the methods described in Appendix 9.3.2). These are presented in Table 5 under the $ra - rc$ heading, with the forest cover change rate in the treated area provided (ra) as a benchmark. In cases where multiple estimates of the same effect were reported (e.g., Arriagada et al. 2012 provide four different estimates of the effects of PSA on forest cover over 1997-2005), we use the mean of the estimates and mean of the associated standard errors (see fn. 8). Section 11.9 of the appendix lists all the effects that we used to compute these mean effects.

Figure 4 provides forest plots of these standardized effects on forest cover change rates. The top panel in Figure 4 shows effects on annual forest cover change rates as measured on forest cover change attributable solely to deforestation (such effects do not account for non-forested areas becoming forested). The bottom panel shows effects on annual forest cover change rates as measured on forest cover change attributable to either deforestation or forest growth on previously non-forested parcels. A random effects mean estimate for the top panel suggests that PES programs have, on average, caused the annual forest cover change rate to be about 0.21 percentage points higher (s.e.=0.09, 95% CI: [0.03, 0.39])¹². In

¹² The random effects estimates were computed in R using the 'metafor' package (Viechtbauer, 2010). The fit was produced using the empirical Bayes procedure, the confidence interval for the random effects mean applies the small sample adjustment of Knapp and Hartung (2003), and the predictive interval uses a t-distribution with degrees of freedom equal to the number of studies minus 1, following Higgins et al. (2009).

other words, PES has tended to reduce the annual rate of deforestation by 0.21 percentage points. The statistics shown at the bottom left of the forest plot are measures of effect heterogeneity generated from the random effects fit. The low (essentially 0) estimate for the between variance (τ^2) and percentage of variability due to between study heterogeneity (I^2) suggest that the effects are highly similar across these studies.

Looking at the bottom forest plot, we see that effect sizes tend to be larger when we look at forest cover change attributable to either deforestation or forest expansion (bottom panel of Figure 4). Estimated effects on annual forest cover change rates ranged from 0.50 percentage points (s.e.=0.20, 95% CI: [0.11, 0.89]) for the Arriagada et al. (2011) study in Costa Rica to 1.6 percentage points (s.e.=0.80, 95% CI: [0.03, 3.17]) for the Honey-Roses et al. (2011) study in Mexico. The Scullion et al. (2011) study reports an effect of 10 percentage points, an outlier in its magnitude (evident in Figure 3). The estimates from Scullion et al. (2011) were not accompanied by standard error estimates, and so we can only report point estimates¹³. Based on our concerns noted above about potential selection biases and spill-over problems, we believe that these estimates likely overstate the true effects of PES on forest cover change.

On the one hand, these larger effects make perfect sense mathematically: the deforestation outcome metric is a truncated measure (at best you can only have zero deforestation) while the forest cover change metric can vary more freely (at best you could have large amounts of forest gain, in principle). At the same time, there is substantive importance of this difference in effect magnitudes. It suggests that evaluations of PES effects ought to take into account not just protection of existing forest but the possibility that PES could contribute to growth or regrowth of forest. This is consistent with the proposition by Daniels et al. (2010) that for Costa Rica's PSA at least, PES was more likely to tip farmers into allowing regrowth on non-forested parcels as opposed to inducing farmers to desist from clearing parcels.

The last effect displayed in the bottom forest plot in Figure 4 shows the estimate for the effects of MBCF on the forest disturbance in Mexico, from Honey-Roses et al. (2011). These authors measure forest disturbance in terms of whether a parcel is covered by forest with at least 70 per cent canopy cover, while deforestation is measured in terms of whether a parcel is covered by forest with any detectable canopy cover. By construction, forest cover disturbance occurs at a higher rate than forest cover change per se, and the point estimate for the effect on the annual change rate is much larger (1.6 percentage points rather than 0.3, see Table 5), although as Figure 4 makes clear, the estimate is quite imprecise. In section 9.4 of the appendix, we present graphs that show the implications of these effects on rates of change for forest cover trajectories.

Many of the study authors (e.g., Robalino and Pfaff [2013], Robalino et al. [2008], and Alix-Garcia et al. [2012]) raise the issue of the inefficiency of PES programs when examined in terms of their forest conservation impacts. There are two parts to the inefficiency equation.

¹³ In section 9.3.1 of the appendix, we indicate that for other studies, we impute a standard error corresponding to a p-value of 0.5 in cases where no standard errors are reported. For this case, given the extreme magnitude of the estimates effects, the resulting confidence intervals would span almost the entire visible range in the graphs, making them completely uninformative. Given the outlier nature of the effect estimates from Scullion et al. and our concerns about its methodological quality, we choose not to proceed with imputing a standard error and otherwise producing synthetic estimates based on the results of this study.

First are the high fixed costs of setting up and managing such programs, given the need to demarcate and measure ex ante forest cover in parcels and then for reliable surveillance methods to monitor compliance. This is an issue that we discuss below in our discussion of the importance of local administrative capacity (see also Honey-Roses et al. [2009] for a detailed discussion). The second problem arises from the fact that by the evidence in the studies we reviewed, the 'additionality' achieved by PES programs has been in the neighborhood of 0.2 percentage point reductions in the annual deforestation rate. This means that after 10 years of programming, we would expect about 98 per cent of forested lands retained to have been retained anyway even were there no PES program in place. As such, the vast majority of parcels on which payments are issued are made 'for nothing' from an ex post perspective. In their qualitative research on reasons for participation in Costa Rica's PSA, Arriagada et al. (2009) obtained the impression that farmers tended mostly to enrol forest areas that they had no intention to deforest.

Arriagada et al. (2012) attempted to characterize what the effects of a more targeted program would be. They focused on Costa Rica's PSA programming in the Sarapiquí region, which was noted for high rates of deforestation. Also, the implementation of PSA in Sarapiquí was facilitated by an NGO that sought to improve targeting. Even there, the estimated effect on the forest cover change rate was a boost of about 0.7 to 1.7 percentage points with a mean effect of 1.2 percentage points (s.e.=0.40, 95% CI: [0.42, 1.98]), implying circa 12 percentage points more land covered in forest after ten years than would have been the case otherwise. This means that after 10 years payments on about 88 per cent of land covered by the PES program would have been for naught when strictly considering forest conservation impact.

As Daniels et al. (2010) point out, it is difficult to know whether the results from the Arriagada et al. (2012) study really represent a 'best case' scenario, given that estimation of the impact of PSA in Costa Rica is confounded generally by the fact that the 1996 law establishing the program also declared blanket regulation on the clearing of natural forests. As Rojas and Aylward (2003) also point out:

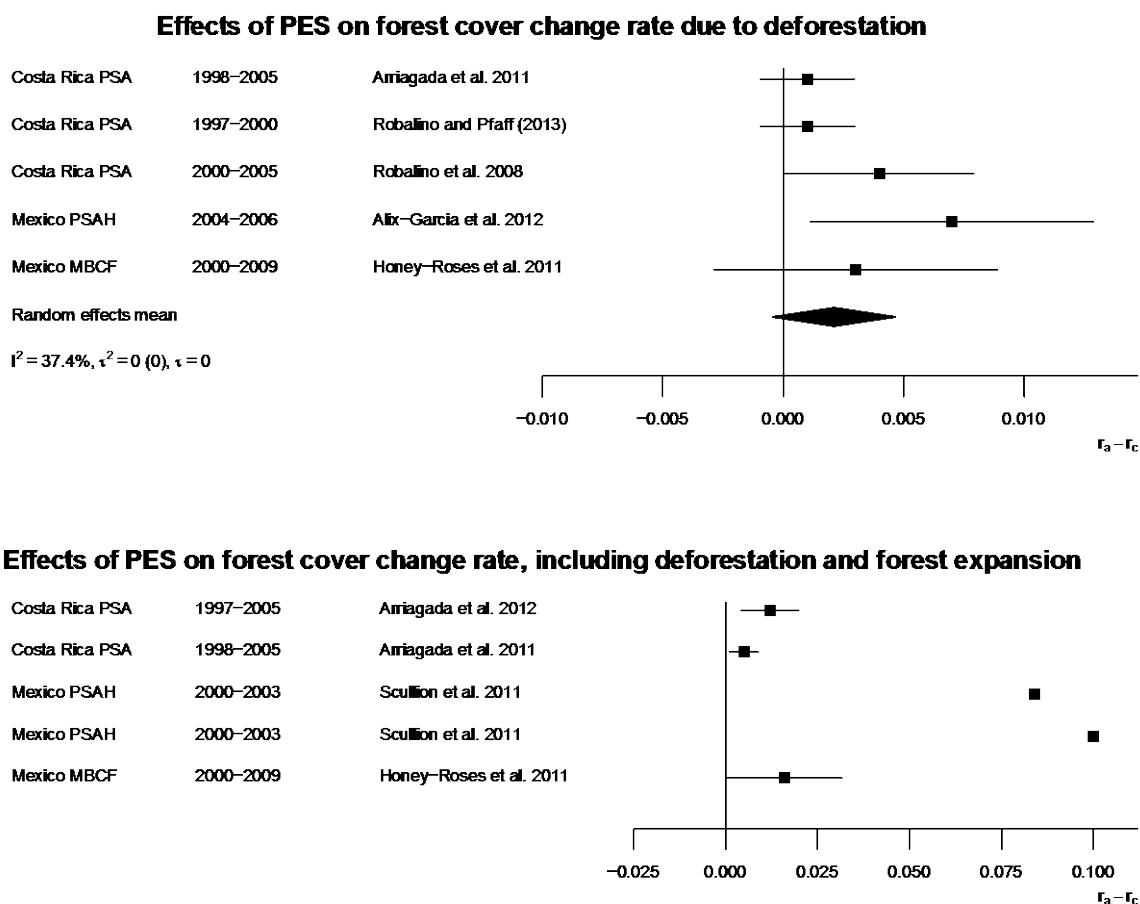
In the Costa Rican case, existing legislation complicates the issue [of studying the impact of the PSA program on forest cover change] as the new [1996] Forestry Law effectively expropriated land use rights on private land by forbidding any change in land use on lands with forest cover. As a result the PES are frequently regarded as a compensatory payment for this expropriation rather than an incentive or compensation per se. (Rojas and Aylward 2003: 94)

The question of the feasibility of a more efficient PES programs remains a question that is open and needs further research. Work on variable pricing and PES auction schemes are one approach, although it is not yet clear that such complex contracting arrangements are practical in developing country contexts (Cason and Gangadharan, 2004; Jack et al., 2008). Munoz-Piña et al. (2008) propose that such mechanisms were explicitly rejected for their complexity in the design of Mexico's PSAH.

Table 5: Studies on PES and forest conservation outcomes

Program	Studies	Treatment Group	Time Period	Counterfactual Approximation	Forest Conservation Outcome	Forest Conservation Effect (standard error in parentheses)	ra-rc (standard error in parentheses)	ra
Costa Rica PSA	Amiagada et al. 2012	1997-2005 PSA conservation contract farms in Sarapiquí	1997-2005	(mean of estimates)	hectares of forest cover	8.730 (3.02)	0.012 (0.004)	0.010
Costa Rica PSA	Amiagada et al. 2011	1998-2005 census tracts covered by PSA conservation contracts and on the common support with matched non-PSA tracts	1998-2005	nearest-neighbor matched census tracts without PSA (best covariate balance estimate) with multiple regression adjustment	hectares of forest deforested	-7.075 (3.91)	0.001 (0.001)	-0.011
Costa Rica PSA	Amiagada et al. 2011	1998-2005 census tracts covered by PSA conservation contracts and on the common support with matched non-PSA tracts	1998-2005	nearest-neighbor matched census tracts without PSA (best covariate balance estimate) with multiple regression adjustment	hectares of forest cover	29.443 (11.43)	0.005 (0.002)	0.010
Costa Rica PSA	Robalino and Pfaff (2013)	1997-2000 1kmX1km forested PSA conservation contracts pixels	1997-2000	(mean of estimates)	percent of forest area deforested	-0.335 (0.18)	0.001 (0.001)	-0.002
Costa Rica PSA	Robalino et al. 2008	2000-2005 1kmX1km forested PSA conservation contracts pixels	2000-2005	(mean of estimates)	percent of forest area deforested	-1.895 (0.7)	0.004 (0.002)	-0.011
Costa Rica PSA	Sanchez-Azofeifa et al. 2007	1997-2000 grid cells covered by PSA conservation contract	1997-2000	regression adjusted grid cells without PSA	probability that area was deforested	-0.002 (0.01)	0.001 (0.002)	-0.003
Costa Rica RISEMP	Garbach et al. 2012	2002-2008 small/medium cattle farm properties in Esparza covered by PSA	2002-2008	convenience sample of local non-PES farmers	number of silvopastoral practices adopted	0.175 (0.37)		
Costa Rica RISEMP	Garbach et al. 2012	2002-2008 small/medium cattle farm properties in Esparza covered by PSA and technical assistance	2002-2008	convenience sample of local non-PES farmers	number of silvopastoral practices adopted	0.791 (0.21)		
Mexico PSAH	Alix-García et al. 2012	2004-2006 PSAH properties	2004-2006	(mean of estimates)	percent of forest area deforested	-1.287 (0.38)	0.007 (0.003)	-0.005
Mexico PSAH	Scullion et al. 2011	2000-2003 PSAH or FIDECOAGUA properties	2000-2003	local properties not under PSAH or FIDECOAGUA	hectares of pine oak forest cover	57 (na)	0.084 (na)	-0.004
Mexico PSAH	Scullion et al. 2011	2000-2003 PSAH or FIDECOAGUA properties	2000-2003	local properties not under PSAH or FIDECOAGUA	hectares of cloud forest cover	264 (na)	0.100 (na)	-0.031
Mexico MBCF	Honey-Roses et al. 2011	2000-2009 MBCF properties in Monarch Butterfly Biosphere Reserve area	2000-2009	spatial matched properties without MBCF	percent of forest area still covered by forest	2.6 (2.1)	0.003 (0.003)	-0.010
Mexico MBCF	Honey-Roses et al. 2011	2000-2009 MBCF properties in Monarch Butterfly Biosphere Reserve area	2000-2009	spatial matched properties without MBCF	percent of forest area undisturbed	11.6 (5)	0.016 (0.008)	-0.018

Figure 4: Estimates of the effect of PES on forest cover change rates due to deforestation (top) and due to either deforestation or forest expansion (bottom).



The small black squares show the point estimates and the horizontal lines running through the squares show 95% confidence intervals. Effects are measured in terms of changes to annualized forest cover change rates (see appendix section 9.3.2 for details). Effect estimates from Scullion et al. (2011, bottom) were not accompanied by standard error estimates in the original study. The text beside each estimate shows the program, timeframe of the program, and the study. The Scullion estimates are for pine oak forest (top) and cloud forest (bottom). The black diamond on the top forest plot displays the random effects synthetic mean estimate of the effect sizes (see fn. 12). Even though studies in the top plot overlap temporally and by country, they are estimates of program effects from different programs in different regions, and so no adjustment for dependent effect sizes was deemed necessary in producing this synthetic mean estimate. No such synthetic mean was produced for the bottom plot because the Scullion study does not provide uncertainty estimates and the Honey-Roses et al. study does not measure outcomes on a scale that is directly comparable to the other studies in the plot.

6.3. Effects of PES on poverty

The evidence base on the effects of PES on welfare, and in particular poverty, is extremely limited. We identified only two quantitative studies reporting effects on human welfare outcomes, covering China's Sloping Land Conversion Program over three years from 2001 to 2004 (Liu et al. 2010) and Mozambique's NCCL program over three years from 2003 to

2006 (Hegde and Bull, 2011). These studies are described in Table 6 below, which also lists the studies' estimates of effects on income. Liu et al. (2010) estimate that the SLCP boosted participating households' incomes by about 14 per cent on average (s.e.=3.42, 95% CI: [7.3, 20.7]). Hegde and Bull (2011) estimate that the NCCL boosted participating households' incomes by about 4 per cent on average (s.e.=1.55, 95% CI: [0.96, 7.04]). Hegde and Bull (2011) also indicate that payments amounted to about 10 per cent of average household income, in which case the fact that the effects on household incomes was only 4 per cent is indicative of opportunities forgone, although they do not provide further clarity on what those may have been.

However, these average effects do not tell us how these programs affect poor households. First, we need to consider the extent to which poor households are really able to participate. For PES to contribute to poverty reduction, poorer households must be able to participate at high rates. But participation in PES programs is typically more difficult for poor households than wealthier households. Pagiola (2008), for example, discusses how Costa Rica's PSA program was structurally biased away from benefitting the poor by nature of the program requirements. For example, the need for land title, proof of tax compliance, absence of debts or fines, production of a notarized land map, and other high transaction costs may have precluded many poor from participating in the first place.

In qualitative and survey interviews with Costa Rican farmers, Arriagada et al. (2009) found that among those who were qualified to participate in PSA, based on land holdings, the primary reason for non-participation was because they did not understand the program. Zbinden and Lee (2005: 269) provide the following interpretation: 'Establishing a PSA contract requires considerable knowledge and the ability to manage administrative tasks. Less education (and presumably often poorer) farmers appear, on average, to be less likely to possess the skills needed to take equal advantage of the forest incentive program made available by the government'. As such, Pagiola (2008: 721) concludes that 'the bulk of program benefits tend to go to larger and relatively better off farmers', a reflection of how Costa Rica's PSA was designed, ostensibly, to prioritize conservation impact.

Alix-Garcia et al. (2012: 619) propose that similar constraints apply in Mexico's PSAH. Focusing on cost-based decisions to participate, their theoretical analysis shows that if PES 'just barely compensates for the opportunity cost of forest production, it will lead to an increase in the amount of land in agriculture [rather than PES] for credit-constrained households'. Access to credit would be required to sacrifice immediate conversion of forested land in anticipation of future benefits. Given that the poor are typically credit-constrained, they are therefore less likely to take up PES contracts of the type that PSAH offered, which were designed to hew closely to opportunity cost projections (Munoz-Piña et al. 2008). For PSAH, about 31 per cent of PSAH payments beneficiaries were poor households (Munoz-Piña et al. 2008). But this figure needs to account for the fact that many such households were benefiting as part of collective contracting through common-property forest user groups known as ejidos (about 70% of the natural forest in Mexico that was eligible for PSAH was on common property [Munoz-Piña et al. 2008]).

Second, one needs to account for the possibility that the effects for poor households may differ substantially from other types of households. Poor households may have less land to

commit to such programs, reducing the potential for benefit in absolute terms¹⁴. Or, it could be that poor households are less able to translate whatever income or freed-up labor the program offers to other income-generating activities that do not require the use of land¹⁵. When weighed against opportunities forgone, this could mean that net benefits for poor households are small.

Hegde and Bull (2011) provide analysis of this issue for the NCCL program in Mozambique. When they focused on only the poor households in their sample, they found no statistically significant benefits to poor households. Looking at consumption expenditure for example, they produced multiple estimates of the effect using different matching methods. The mean of their effect estimates for the whole sample was MTS 70,039/capita (s.e.= 28,778, 95% CI: [13634, 126444]). For poor households, the mean of their effect estimates was insignificant (point estimate of MTS 13,933/capita, s.e.=23,425, 95% CI: [-31980, 59846]). For cash income, the mean of their effect estimates for the whole sample was MTS 278,750/capita (s.e. = 108,043, 95% CI: [66986, 490514]), implying about a 4 per cent gain (s.e.=1.55 percentage points, 95% CI: [0.96, 7.04]) given average income levels in the study area. For poor households, the mean of their effect estimates was again insignificant (point estimate MTS 103,256/capita, s.e. = 154,113, 95% CI: [-198805, 405317]). Hegde and Bull's (2011) results suggest that poor households did not benefit as much as other households, in absolute terms, as a result of their participation (although we should note that they did still benefit to some extent).

Hegde and Bull's (2011) analysis does not try to disentangle the precise reasons for the difference in the poor versus non-poor households, and the reasons for such variable benefits should be investigated in future research. For example it may be that consumption expenditure is a poor measure for the welfare of poorer households in this context given limited access to markets.¹⁶ A better measure may be savings and the maintenance of 'rainy day funds' to cover adverse economic or health shocks. For PSAH in Mexico, Munoz-Piña et al. (2008: 6) find that 'few ejido members aside from those with directive or representation functions know the conditions of the [PSAH] contract, even in ejidos that distributed payments among members'. While this does not guarantee that those in prominent positions in the ejido 'captured' the system for their benefit, it does make one concerned about this possibility. Additional studies should investigate this concern.

Third, a proper analysis of the effects of PES on poverty cannot simply look at the effects of participation in PES programs, especially since participation per se is likely to be difficult for many poor households. Rather, one needs to consider how PES programs implemented in an area may also affect the welfare of poor households in the program area even if they are non-participating households. If PES reduces the demand for labor associated with logging

14 As a reviewer rightly pointed out, it may be that landless households may still participate in PES programs that provide payments for people to desist from deforesting lands that are not theirs---e.g., forests in protected areas. Of course, such a program would provide no additionality above what the protected area is intended to conserve.

15 Of course, if both the poor and wealthy have excess land, then it would be the poor who would be less likely to make full use of it for lack of access to inputs. In that case, it could be that the poor stand to gain the most from PES, although it is not clear to what extent such conditions of excess land among the poor are relevant. We thank a reviewer for this point.

16 We thank a reviewer for this raising this point.

or agriculture, then poor labor-supplying households may suffer. We did not identify any studies looking into this question with respect to PES, unfortunately, and so we propose this as a priority for further research. Relevant references for such research are studies by Robalino and Villalobos-Fiatt (2010) and Sims (2010) on the effects of conservation parks and protected areas on welfare outcomes.

Table 6: Studies on PES and poverty outcomes

Program	Studies	Treatment Group	Time Period	Counterfactual Approximation	Welfare Outcome	Welfare Effect (standard error in parentheses)	Percentage Effect (standard error in parentheses)
China SCLP	Liu et al. 2010	2001-2004 SLCP households	2001-2004	regression adjusted households without SLCP	Natural log of total household income	0.13 (0.03)	14% (3.42)
Mozambique NCCL	Hegde and Bull 2011	2003-2006 NCCL participant households in project region	2003-2006	(mean of estimates)	Consumption expenditure per capita in MTS	70039 (28778)	*
		(poor households)		(mean of estimates)	Consumption expenditure per capita in MTS	13933 (23425)	*
Mozambique NCCL	Hegde and Bull 2011	2003-2006 NCCL participant households in project region	2003-2006	(mean of estimates)	Household cash income per capita in MTS	278750 (108043)	4% (1.55)
		(poor households)		(mean of estimates)	Household cash income per capita in MTS	103256 (154113)	18% (38.13)

6.4. Intersection of poverty and deforestation impact

Our theoretical discussion above proposed that the conservation impact of a PES program might be tied to its poverty impact. In order to address this possibility, one would need studies that evaluate both poverty and conservation outcomes jointly. Unfortunately, no such studies were identified. Surprisingly, there was no overlap in the programs covered by the quantitative studies on conservation and poverty, respectively, and so even general comparisons across studies are impossible. That being the case, we cannot address the mediation hypothesis.

Some studies reported on how prevailing conditions of poverty might moderate the impact of PES programs. For example, Alix-Garcia et al.'s (2012) theoretical analysis of PES proposes that, conditional on a fixed payment schedule, poorer households will be all the more sensitive to opportunity costs if they are more credit-constrained and therefore less able to smooth over forgone immediate production in favour of future PES payments. For this reason, we might expect parcels put under PES contract by poor households to carry the lowest opportunity costs to conservation, or in other words, to be the least likely to be subject to deforestation pressure among parcels included in a PES program.

Consistent with this reasoning, Kerr et al. (2004) find that poorer farmers in Costa Rica are the least responsive among farmers in general to changes in the productivity returns to land, suggesting that credit constraints and other factors that limit the ability of poor farmers to switch across different types of production. In a more direct empirical assessment, Alix-Garcia et al. (2012) find that the conservation impact in poorer municipalities (measured using a municipality vulnerability index) is substantially lower. While PES impacts in the richest areas were twice that of the overall sample, in the poorest areas the impact on avoided deforestation was approximately zero. This has important implications for debates about the extent to which welfare considerations, and particular poverty impact, should be incorporated into the design of PES programs. At least with respect to Mexico's PSAH, more targeting of poor areas would only have served to reduce the program's conservation impact.

6.5. ROLE of institutional and social context

The qualitative studies contain some insights on how the institutional and social context bears on the design and performance of PES programs. In some cases, we obtained specific insights relevant to our hypotheses about how local administrative capacity, corruption, democratic accountability, and inequality moderate the effectiveness of each type of program. Qualitative analyses of Mexico's PSAH by McAfee and Shapiro (2010) and Munoz-Piña et al. (2008) describe how the institutional and social context forced the design of PSAH to deviate from what one might consider the most economically optimal. These analyses touch on all three of the obstacles that Wunder et al. (2008) highlight as crucial challenges to efficient PES program design: (1) fairness and political constraints, (2) corruption and rent seeking through the program, and (3) capacity and knowledge limitations. These three factors line up very well with our proposed moderating factors. Analyses of PSA in Costa Rica were much less adamant about how PSA deviated from what was perceived at the time as an efficient PES program, perhaps indicative of the less contested nature of the PSA policy-formulation process in Costa Rica relative to contestation around PSAH in Mexico. We summarize the findings from these accounts in the sections that follow.

6.5.1. Inequality

Our theoretical discussion proposes that high levels of inequality in forest areas may undermine the effectiveness of PES programs. The qualitative data suggests the Mexican PSAH program was clearly sensitive to concerns of inequality. Wunder et al. (2008: 850) point out that 'efforts to spread payments 'fairly' throughout the country meant that a substantial share of funding went to area at little risk of deforestation and/or with limited or no threats to water supplies'. Inequality was such a large issue that the program focused more on keeping payments equal than on conserving the most endangered areas. McAfee and Shapiro (2010: 3) corroborate this, finding that 'involvement of federal agencies and rural activists shifted the program's emphasis toward poverty alleviation', and officials involved in the program took this to imply that the conservation impact would be lessened.

6.5.2. Capacity

Our theoretical discussion proposed that the level of local state capacity positively moderates the conservation impact of PES programs. We cannot evaluate this hypothesis quantitatively given the low number of studies. Also, the impact estimates do not vary

enough for us to rank clearly the success of the programs and implementation periods. Nonetheless, qualitative accounts provide some useful insights. For PES, Honey-Roses et al. (2009) discuss the technical challenges of monitoring PES participants for compliance. Their review of current standards for using remote sensing and field surveying to monitor forest cover change leads them to conclude that 'current PES and PES-like schemes are underestimating the land-use changes and overpaying non-compliant participants' (Honey-Roses et al. 2009: 126). As Munoz-Piña et al. (2008) put it:

The PSAH program reported no deforestation in participating areas. The claim of 100 per cent compliance is difficult to believe in the Mexican context, especially when the seriousness of cancellation of payments has not yet been experienced by any forest owner. [National Ecological Institute] researchers believe that the current low-resolution monitoring method is responsible for the over-enthusiastic results (Munoz-Piña et al. 2008: 8).

Better monitoring to boost PES impact would require even higher levels of administrative capacity than have been applied thus far.

6.5.3. Corruption

Our theoretical analysis suggests that corruption in government negatively moderates conservation impact of PES programs. Some of the qualitative studies provide insights, although again no direct quantitative test is possible from the available data. At its core, PES is a payment distribution mechanism. As such, the potential for corruption is high. Corrupt institutions may be more likely to siphon off payments and prevent them from being delivered to service providers. Corrupt officials may be less capable of or interested in enforcing conservation regulations. Wunder et al. (2008) described how payments in Mexico's PSAH program were used to finance things that have little to do with conservation: 'Many side objectives in Mexico's PSAH program, for example, were added after the program was created, either to placate politically powerful groups or to address other government objectives for which funds were insufficient' (Wunder et al. 2008: 849).

Muñoz-Piña et al. (2008) also describe corruption in the creation of the Mexican PSAH program. During negotiations, landowner organizations used their political clout to fight for—and receive—significantly higher PES payments than what national program staff considered to be appropriate based on an opportunity cost assessment.

6.5.4. Democratic Accountability

The assumption that forest-edge communities place especially high values on conservation needs to be scrutinized. McAfee and Shapiro (2010) found that forest-edge indigenous groups in Mexico were much less supportive of programs that enforced pure conservation than were central government officials charged with resource management policies. In the formulation of Mexico's PSAH program, such communities called for sustainable forest agriculture to be admitted as an activity that qualified for PES payments, rather than only admitting 'no-touch' conservation. Adapting the program to these demands may result in higher levels of forest disturbance. At the same time, it is necessary to appreciate that the indigenous committees were not calling for clear-cutting or full-scale conversion into plantation land. Conservation programs may need to allow for sustainable forest use to achieve forest-edge community support.

In addition to fostering a feeling of participatory accountability in forest dwelling communities, transparent democratic processes also increase the likelihood that forest councils will maintain records of their meetings, finances, and decisions. According to Agrawal (2001), this makes it easier for central government officials to monitor forest management and know when additional resources are needed and where.

6.6. Variation in the Program Design Features

We also documented variation in design features of the PES programs covered by the eligible studies. The number of studies is too few to assess rigorously how such design features may affect program impact. Nonetheless, we can point out some patterns. Table 7 shows program design features for the PES programs included in the synthesis. In all cases for which we could retrieve details, some form of conditionality was in fact applied. For Costa Rica and Mexico, we have complete information on scale of payments. When compared to prevailing income levels, these programs differ markedly: while the contracts in Costa Rica pay on the order of 6-9 per cent of average national income per hectare, the Mexico contracts pay at a rate of less than 1 per cent of average national income per hectare. And yet, the estimated forest conservation effects for the Mexico programs were no smaller, and even appear to be larger, than for the Costa Rica program. This is not what one would expect to see, although our use of national income averages could obscure how the scale of payments relates to those who were actually targeted by the program (e.g. it is conceivable that the Costa Rica program was targeting land owners whose incomes and opportunity costs were an order of magnitude higher than those targeted by the programs in Mexico).

Table 7: Program design features for PES programs

Country	Program	Article(s)	Length	Required Activities	Payment Amounts	Avg. Annual Income	Share of Income	Conditionality
China	Sloping Land Conversion Program (SLCP)	Liu et al. 2010	9 years	sloping/desertified cropland is converted to either ecological or economic forest	subsidies provided in seedlings, grain, and cash (specific amounts not included in article)	7716.67 Yuan (in 2004, study sample)	n/a	n/a
Costa Rica	Pagos por Servicios Ambientales (PSA)	Arriagada et al. 2012 Arriagada et al. 2011 Pfaff et al. 2008 Robalino et al. 2008 Sanchez-Azofeifa et al. 2007	7-8 years	conservation reforestation forest management	\$210/ha/year - conservation \$537/ha total - reforestation \$327/ha total - forest management	\$3500 (ca. 2000, World Bank GNI per capita)	6%-9% of average annual income per ha	site visits conducted to ensure compliance
Costa Rica	Regional Integrated Silvopastoral Approaches to Ecosystem Management (RISEMP)	Garbach et al. 2012	6 years	improved pasture simple live fence pasture trees - low density forage banks pasture trees - high density complex live fence riparian forest	based on composite index of ha covered and actions taken (specific amounts not included in article)	\$3500 (ca. 2000, World Bank GNI per capita)	n/a	n/a
Mexico	Pago por Servicios Ambientales Hidrológicos (PSAH)	Alix-Garcia et al. 2012 Scullion et al. 2011	5 years	maintain existing forest cover	\$36/ha/year - cloud forest \$27/ha/year - pine/oak forest	\$8500 (ca. 2009, World Bank GNI per capita)	<1% of average annual income per ha, <3% of annual income for total received by households in study sample (as reported by authors)	forest cover verified via satellite, deforested areas removed from study and payments adjusted accordingly
Mexico	Monarch Butterfly Conservation Fund (MBCF)	Honey-Roses et al. 2011	9 years	communities must forego logging rights due to establishment of protected areas	\$18/m ³ of forfeited timber \$10-12/ha of conserved forest	\$8500 (ca. 2009, World Bank GNI per capita)	<1% of average annual income per ha	payments withheld in cases of non-compliance
Mozambique	Nhambita Community Carbon Livelihoods (NCCL)	Hegde and Bull 2011	25 years	plant trees and manage them	money (specific amount not included in article) split into 7 annual payments: 30% in year 1 12% in years 2-6 10% in year 7	MTS 7,184,881 (PES participants) MTS 5,012,907 (non-participants)	n/a	yes

7. Authors' Conclusions

7.1. Summary of findings with respect to our hypotheses

Our analysis sought to test two main hypotheses and then a set of mediator and moderator hypotheses:

Main hypotheses:

1. H1: PES reduce deforestation rates.
2. H2: PES have non-negative impact on local poverty levels.

Mediator hypotheses

3. H3: The more a PES program functions to relieve poverty, the stronger will be its impact on reducing deforestation.
4. H4: PES deforestation reduction impact is negatively moderated by the level of local inequality.

Moderator hypotheses

5. H5: PES deforestation reduction impacts are positively moderated by the level of local administrative and enforcement capacity.
6. H6: PES deforestation reduction impacts are negatively moderated by the level of corruption in government.

Limitations of the evidence base preclude definitive tests of any of these hypotheses. With respect to hypotheses 1, we do find that PES reduce deforestation rates on average. The effect estimates suggest the impact is modest and seems to come with extremely high levels of inefficiency. For hypotheses 2, we cannot say that the evidence indicates non-negative effects on poverty for PES. This is a troubling finding, but it is based on only a handful of cases and therefore deserves much more empirical attention. We were unable to assess hypotheses 3, although we find that areas of higher levels of poverty tend to be associated with poorer conservation performance. We found qualitative evidence in support of hypotheses 4, 5 and 6, suggesting that the contextual conditions of inequality, limited local administrative and enforcement capacity, and corruption may undermine the effectiveness of PES programs. However, in the absence of clear tests, these findings remain highly uncertain.

7.2. Implications for policy: Elusive win-win

Our review sought to address the fundamental issues of the extent to which conservation and poverty reduction goals conflict, how different conservation strategies fare in terms of such trade-offs, and the scope for 'win-win' strategies that generate both significant conservation and poverty reduction benefits. We outlined two sides of the argument about the extent to which conservation and poverty reduction goals ought to be married to each other. After reviewing the evidence, we are largely in agreement with the type of 'guarded pessimism' reflected by Wunder (2001; 2013) and the notion that PES probably offers mostly 'win-settle' solutions (Wunder, 2013), at least when it comes to strategies that have been pursued to date. Pagiola (2004) asks, 'can payments for environmental services help reduce poverty?' Our review has tried to take the enquiry a step further and asked, based on the

accumulated evidence, should PES programs have poverty reduction as a part of joint goal with conservation? A pragmatic logic for doing so is that poverty alleviation benefits may help to motivate better performance in providing forest conservation services. Hope for a win-win is rooted further in the idea that property rights to forest areas are 'often the only capital of the poor who have no money or political voice' (Arriagada et al., 2009: 344), and that PES programs allow for the conversion of such property to income.

The available data provide scant evidence for addressing this question, but based on what we have seen on patterns of participation in PES programs and their welfare impacts, there is no basis to claim that PES programs are 'pro-poor', in fact, the opposite may be true. Furthermore, poverty goals and conservation goals do appear to conflict in a manner that advises against setting poverty reduction as a goal for PES programs, at least for PES programs that set natural forest conservation as the primary objective.

Of course, strategies that have been evaluated to date may be limited in terms of the lessons that they offer for potential synergy in conservation and poverty-reduction programming. If so, then to the extent that these twin goals need to be pursued jointly, new ideas and program concepts must be developed. The PES programs that we included in our synthesis shared some important features that limit the generality of our conclusions: (1) they diverged from the 'ideal type' PES program insofar as buyer entities (usually, governments or NGOs) were not direct consumers of the environmental services and (2) the goal was most often conservation of natural forest rather than sustainable forest use. In this way, our conclusions cannot be taken as a summary judgment on the entire class of potential PES programs, and particularly not for programs that try either to maximize the role of 'service consumer' feedback or, and this may be more important from a poverty perspective, to promote behaviors such as sustainable forest use rather than strict conservation per se. The importance of the latter scope condition is made especially clear by McAfee and Shapiro (2010) in their analysis of the politics behind Mexico's PSAH.

As implemented, the evidence that we review indicates that government- and NGO-administered forest conservation PES programs have also been rather inefficient in producing forest conservation services. Evidence from Alix-Garcia et al. (2012) suggests that orienting Mexico's PSAH even more toward poorer communities in Mexico would have heightened this inefficiency. Hegde and Bull (2011) found that in Mozambique, the ability of households to capitalize on PES income to improve consumption and overall welfare (e.g., by using PES income to finance intensification of other productive activities) was least likely to happen among poor households. As a result, PES income barely substituted for opportunities forgone, possibly even causing a net reduction in welfare.

Without further evidence to the contrary, there is no evidence of the type of 'win-win' that would motivate combining poverty reduction with conservation goals in PES programs. That is not to say that PES programs should avoid seeking to mitigate any harm introduced. It is to highlight a profound tension in the idea that PES ought to have poverty reduction goals, rather than poverty mitigation goals, as a priority.

The lack of an apparent win-win means that the costs from inefficiencies of targeting poor areas for PES are unlikely to be offset by sizable benefits in terms of enhancing welfare for the poor. A first order issue for policy makers working with PES programs is to address the extreme inefficiency when it comes to the amount of payments issued that are unlikely to make any difference in terms of environmental impacts. That is, targeting PES programming

with the first order objective of maximizing environmental return on investment is the main priority for the next generation of PES. If, for reasons having to do with ecological conditions, the areas targeted for PES happen to be areas where poor residents are concentrated, the evidence suggests that this should not be a reason to celebrate, but rather a reason to consider the need for additional, complementary resources to be provided, presumably based on the logic of credit-constraints developed by Alix-Garcia et al. (2012). Such complementary resources would seem to be necessary for a PES program operating in a poor area to have a good chance of succeeding in terms of conservation impact and welfare impact.

7.3. implications for research: Need for more rigorous research across contexts

Our final conclusion re-emphasizes the poor state of the evidence base for conservation programming. Much advanced scientific effort and extensive investment has gone into measuring forest conditions around the world. Relative to that, the evidence base on the ex post performance of PES programs is limited in size and methodologically weak. Composed as it is of a few quasi-experimental studies of varying quality, the evidence base provides a very shaky foundation, likely tainted by selection biases, for informing environmental and development policy making.

As far as we know, there are no completed randomized controlled studies despite the fact that such would seem to be quite feasible. Feasibility of field experimental studies for PES is apparent from the fact that the few high quality quasi-experimental studies that we did review constructed approximations to the treatment group counterfactual using local non-PES properties. One study by Garbach et al. (2012) established a perfect opportunity for a field experimental study by randomly selecting farmers to participate in the RISEMP pilot. But they failed to follow through and track outcomes among the 'control' group that was constructed through this random selection process, choosing instead to use a convenience sample of local farmers as a control. The possibility for learning from the pilot was severely undermined as a result.

While field experiments should be the methodological priority, the quasi-experimental studies covered in this review might be replicable for other countries and programs given tools such as Google Earth Engine's high resolution forest cover mapping (Hansen et al. 2013). Thus, there would seem to be ample opportunity to expand the coverage of these sorts of quasi-experimental studies around the world as formative research that might inform more finely targeted field experimental studies.

Future experimental and quasi-experimental studies should assess both the environmental and human welfare outcomes of PES to allow and assessment of potential synergies or trade-offs between different program objectives. Quantitative studies should also collect data on context, implementation and costs.

Moreover, the existing evidence base is limited to a few countries, and excludes vast experience from other parts of the world. We were surprised to find no studies from countries with large PES programs, such as Indonesia or Brazil. Future research should focus on assessing the effects of PES across a diversity of contexts, including in particular contexts with high de-forestation rates.

Finally, the results above suggest that priority topics for further research include (i) mechanisms for more efficient contracting and (ii) strategies to boosting conservation performance in poor areas, such as allowing sustainable use (as opposed to only non-use) of forest lands to qualify for payments.

7.4. Limitations and deviations from protocol

Limitations of this study derive from the very few cases that the quantitative evidence base covers. The countries that we cover in this review exclude the major forested areas in the tropics, including the forests of the Amazon Basin, Indonesia, and the Congo Basin.

Details on the deviations from protocol are listed in section 11.5 of the appendix. The key point that we make there is that the very limited extent of the database prevented us from being able to do the type of thorough analysis of factors that moderate the effectiveness of PES programs. Neither were we able to investigate directly how deforestation and poverty alleviation goals interact since we found studies that looked at effects on these outcomes jointly.

7.5. Acknowledgements

We thank the authors of the studies included in this review for their excellent work and the feedback they provided to us on earlier drafts of this review. We received enormously helpful feedback from three anonymous reviewers, the Campbell Collaboration International Development Coordinating Group Editor, Birte Snilstveit, and Methods Editor, Terri Pigott; members of an advisory group that included Michele de Nevers, Ilaria Firmian, Steven Lawry, Carlos Manuel Rodriguez, Jesus Quintana, and David Young; conference participants at the 2013 American Economic Association annual meeting, 2012 3ie Systematic Review Colloquium, and the 2014 Campbell Colloquium; and David Kaimowitz and Sven Wunder. Reema Badrinarayan, Harini Dedhia, Justin Epstein, Shomik Ghosh and Johan Gorr provided excellent research assistance.

References

All references

* indicates inclusion for quantitative synthesis.

indicates inclusion for qualitative, but not quantitative, synthesis.

Abadie, A., Athey, S.A., Imbens, G.W., & Wooldridge, J. (2014). Finite Population Causal Standard Errors. Manuscript, Harvard University, Stanford University, and Michigan State University.

Achard, F, & Hansen, M.C., eds. (2013). Global Forest Monitoring from Earth Observation. Boca Raton, FL: CRC Press.

#Agrawal, A. (2001). The Regulatory Community: Decentralization and the Environment in the Van Panchayats (Forest Councils) of Kumaon, India. *Mountain Research and Development*, 21(3), 208-211.

Agrawal, A., & Angelsen, A. (2009). Using Community Forest Management to Achieve REDD+ Goals. In A. Angelsen, ed., *Realising REDD+: National Strategy and Policy Options*. Bogor: CIFOR.

Agrawal, A., & Benson, C.S. (2011). Common Property Theory and Resource Governance Institutions: Strengthening Explanations of Multiple Outcomes. *Environmental Conservation* 38(2), 199-210.

Akerlof, G. (1982). Labor Contracts as Partial Gift Exchange. *Quarterly Journal of Economics*. 97:543-569.

*Alix-Garcia, J., Shapiro, E.N., & Sims, K.R.E. (2012). Forest Conservation and Slippage: Evidence from Mexico's National Payments for Ecosystem Services Program. *Land Economics*, 88(4), 613-638.

*Andersson, K., & Gibson, C.G. (2007). Decentralized Governance and Environmental Change: local Institutional Moderation of Deforestation in Bolivia. *Journal of Policy Analysis and Management*, 26(1), 99-123.

Angelsen, A. ed. (2009). *Realising REDD+: National Strategy and Policy Options*. Bogor: CIFOR.

Angelsen, A. (2010). Policies for Reduced Deforestation and Their Impact on Agricultural Production. *Proceedings of the National Academy of Sciences*, 107(46), 19639-19644.

Angelsen, A., & Wunder, S. (2003). Exploring the Forest-Poverty Link: Key Concepts, Issues, and Research Implications. *CIFOR Occasional Paper*, 40.

*Arriagada, R.A., Sills, E.O., & Pattanayak, S.K. (2011). Payments for Environmental Services and Their Impact on Forest Transition in Costa Rica. Working Paper.

*Arriagada, R.A., Ferraro, P.J., Sills, E.O., & Pattanayak, S.K. (2012). Do Payments for Environmental Services Affect Forest Cover?: A Farm-Level Evaluation from Costa Rica. *Land Economics*, 88(2), 382-399.

#Arriagada, R.A., Sills, E.O., Pattanayak, S.K., & Ferraro, P.J. (2009). Combining Qualitative and Quantitative Methods to Evaluate Participation in Costa Rica's Program of Payments for Environmental Services. *Journal of Sustainable Forestry*, 28, 343-367.

*Baland, J., Bardhan, P., Das, S., & Mookherjee, D. (2010). Forests to the People: Decentralization and Forest Degradation in the Indian Himalayas. *World Development*, 38(11), 1642-1656.

#Banana, A.Y., Gombya-Ssembajjwe, W., & Bahati, J. (2002). Decentralization of Forestry Resources in Uganda: Realities or Rhetoric? Makerere University. Kampala, Uganda.

Banerjee, A.V., & Duflo, E. (2011). *Poor Economics: A Radical Rethinking of the Way to Fight Global Poverty*. New York, NY: Public Affairs.

#Baral, N., & Heinen, J.T. (2007). Decentralization and people's participation in conservation: a comparative study from the Western Terai of Nepal. *International Journal of Sustainable Development & World Ecology*, 14, 520-531.

Bardhan, P., & Mookherjee, D. 2005. Decentralizing Antipoverty Program Delivery in Developing Countries. *Journal of Public Economics*, 89(4), 675-704.

*Blessings, C., Jumbe, L., & Angelsen, A. (2006). Do the Poor Benefit from Devolution Policies? Evidence from Malawi's Forest Co-Management Program. *Land Economics*, 82(4), 562-581.

Boerner, J., Wunder, S. (2008). Paying for Avoided Deforestation in the Brazilian Amazon: From Cost Assessment to Scheme Design. *International Forestry Review* 10(3), 496-511.

Bowler, D., Buyung-Ali, L., Healey, J.R., Jones, J.P.G., Knight, T., & Pullin, A.S. (2010). The Evidence Base for Community Forest Management for Supplying Global Environmental Benefits and Improving Local Welfare: Systematic Review. *Collaboration for Environmental Evidence Review 08-011 (SR48)*. Environmental Evidence: www.environmentalevidence.org/SR48.html

Burgess, R., Hansen, M., Olken, B., Potapov, P., & Sieber, S. (2012). The Political Economy of Deforestation in the Tropics. *The Quarterly Journal of Economics*, 124(4), 1707-1754.

Campbell Collaboration International Development Coordinating Group (2011). *Systematic Review Protocol and Review Guidelines*. London: Campbell Collaboration.

Casey, K., Glennerster, R., & Miguel, E. (2012). Reshaping institutions: Evidence on aid impacts using a preanalysis plan. *The Quarterly Journal of Economics*, 127(4), 1755-1812.

Cason, T., & Ganghadaran, L. (2004). Auction design for voluntary conservation programs. *American Journal of Agricultural Economics*, 86(5), 1211-1271.

Chomitz, K.M. (2007). *At Loggerheads? Agricultural Expansion, Poverty Reduction, and Environment in the Tropical Forests*. Washington, DC: The World Bank.

Cochrane Effective Practice and Organization of Care Group (2009). *Risk of Bias*. London: Cochrane Collaboration.

- #Cole, R.J. (2010). Social and environmental impacts of payments for environmental services for agroforestry on small-scale farms in southern Costa Rica. *International Journal of Sustainable Development & World Ecology*, 17(3), 208-216.
- Collier, D. (2011). Understanding Process Tracing. *PS: Political Science and Politics*. 44(4):823-830.
- Conte, M., & Kotchen, M. (2010). Explaining the Price of Voluntary Carbon Offsets. *Climate Change Economics*. 1 (2).
- Corbera, E., Gonzalez Soberanis, C., & Brown, K. (2009). Institutional Dimensions of Payments for Ecosystem Services: An Analysis of Mexico's Carbon Forestry Programme. *Ecological Economics* 68, 743-761.
- Daniels, A., Bagstad, K., Esposito, V., Moulaert, A., & Manuel Rodriguez, C. (2010). Understanding the impacts of Costa Rica's PES: Are we asking the right questions? *Ecological Economics*, 69, 2116-2126.
- Deaton, A. (1997). *Analysis of Household Surveys: A Microeconometric Approach to Development Policy*. Baltimore, MD: Johns Hopkins University Press.
- Deveny, A., Nackoney, J., & Purvis, N. (2009). *Forest Carbon Index: The Geography of Forests in Climate Solutions*. Washington DC: Resources for the Future.
- *Edmonds, E.V. (2002). Government-initiated community resource management and local resource extraction from Nepal's forests. *Journal of Development Economics*, 68, 89-115.
- Edwards, R., Hunter, S., Stephan, C., Phamtran, M., Margules, C., & Nurse, M. (2011). *Linking Poverty Alleviation to Ecosystem Service Payments in Asia-Pacific: A Call to Action*. Melbourne: The Foundation for Development Cooperation.
- Ferraro, P.J. (2011). The Future of Payments for Environmental Services. *Conservation Biology* 25(6), 1134-1138.
- Ferraro, P.J., Lawlor, K., Mullan, K.L., & Pattanayak, S.K. (2011). Forest Figures: Ecosystem Services and Policy Evaluation in Developing Countries. *Review of Environmental Economics and Policy* 6(1), 20-44.
- Food and Agricultural Organization. (2010). *Global Forest Resources Assessment 2010: Terms and Definitions*. Rome: United Nations Food and Agricultural Organization.
- Forest Trends, The Katoomba Group, & UNEP. (2008). *Payments for Ecosystem Services: Getting Started*. Nairobi: Forest Trends and the Katoomba Group.
- *Garbach, K., Lubell, M., & DeClerck, F.A.J. (2012). Payment for Ecosystem Services: The roles of positive incentives and information sharing in stimulating adoption of silvopastoral conservation practices. *Agriculture, Ecosystems and Environment*, 156, 27-36.
- Geldman, J., Barnes, M., Coad, L., Craigie, I., Hockings, M., & Burgess, N. (2013). *Effectiveness of Terrestrial Protected Areas In Maintaining Biodiversity and Reducing Habitat Loss: Systematic Review*. Collaboration for Environmental Evidence.
- *Gelo, D., & Koch, S.F. (2012). *Welfare and Common Property Rights Forestry: Evidence from Ethiopian Villages*. Working Paper 277, University of Pretoria.

Gibson, C.G., McKean, M.A., & Ostrom, E., eds. (2000). *People and Forests: Communities, Institutions, and Governance*. Cambridge, MA: MIT Press.

Gregersen, H., El Lakany, H., Karsenty, A., & White, A. (2010). *Does the Opportunity Cost Approach Indicate the Real Cost of REDD+?* Washington, DC: Rights and Resources Initiative.

Grieg-Gran, M. (2008). *The Cost of Avoiding Deforestation*. London: International Institute for Environment and Development.

Gullison, R.E.; Frumhoff, P.C.; Canade, J.G.; Field, C.B.; Nepstad, D.C.; Hayhoe, K.; Avissr, R.; Curran, L.M.; Friedlingstein, P.; Jones, C.D. & Nobre, C. (2007). Tropical Forests and Climate Policy. *Science* 316(5827):985-986.

Hansen, M. C., Potapov, P.V., Moore, R., Hancher, M., Turubanova¹, S.A., Tyukavina¹, A., Thau, D., Stehman, S.V., Goetz, S.J., Loveland, T.R., Kommareddy, A., Egorov, A., Chini, L., Justice, C.O., & Townshend, J.R.G. (2013). High-resolution global maps of 21st-century forest cover change. *Science*, 342(6160), 850-853.

Harris, N., Brown, S., Hagen, S.C., Baccini, A., & Houghton, R. (2012). *Progress Toward a Consensus on Carbon Emissions from Tropical Deforestation*. Arlington, VA, & Falmouth, MA: Winrock International & Woods Hole Research Center.

*Hegde, R., & Bull, G.Q. (2011). Performance of an agro-forestry based Payments-for-Environmental-Services project in Mozambique: A household level analysis. *Ecological Economics*, 71, 122-130.

Higgins, J.P.T., Thompson, S.G., & Spiegelhalter, D.J. (2009). A re-evaluation of random-effects meta-analysis. *Journal of the Royal Statistical Society, Series A*, 172(1), 137-159.

*Honey-Rosés, J., Baylis, K., & Ramírez, M.I. (2011). A Spatially Explicit Estimate of Avoided Forest Loss. *Conservation Biology*, 25(5), 1032-1043.

#Honey-Rosés, J., López-García, J. Rendón-Salinas, E., Peralta-Higuera, A., & Galindo-Leal, C. (2009). To pay or not to pay? Monitoring performance and enforcing conditionality when paying for forest conservation in Mexico. *Environmental Conservation*, 36(2), 120-128.

Intergovernmental Panel on Climate Change. (2007). *Climate Change 2007: Synthesis Report*. Adopted at the Intergovernmental Panel on Climate Change Plenary XXVII, Valencia.

Irawan, S. & Tacconi, L. (2009). Reducing Emissions from Deforestation and Forest Degredation (REDD) and decentralized forest management. *International Forestry Review*, 11(4), 427-438.

Jack, B.K., Kousky, C. & Sims, K.R.E. (2008). Designing payments for ecosystem services: Lessons from previous experiences with incentive-based mechanisms. *Proceedings of the National Academy of Sciences*, 105:9465-9470.

*Jagger, P. (2008). *Forest Incomes After Uganda's Forest Sector Reform: Are the Rural Poor Gaining?* Working Paper No. 92, CAPRI.

Karlan, D., & Appel, J. (2012). *More Than Good Intentions: Improving the Ways the World's Poor Borrow, Save, Farm, and Stay Healthy*. London: Dutton.

- Kauneckis, D., & Andersson, K. (2009). Making Decentralization Work: A Cross-national Examination of Local Governments and Natural Resource Governance in Latin America. *Studies in Comparative Development*, 44, 23-46.
- Kerr, S., Lipper, L., Pfaff, A., Cavatassi, R., Davis, B., Hendy, J., & Sanchez, A. (2004). Will buying tropical forest carbon benefit the poor? Evidence from Costa Rica. *ESA Working Paper No. 04-20*.
- Knapp, G., & Hartung, J. (2003). Improved Tests for a Random Effects Meta-Regression with a Single Covariate. *Statistics in Medicine*, 22(17), 2693-2710.
- Kremen, C., J. Niles, M. Dalton, G. Daily, P. Ehlich, J. Fay, D. Grewal, and R. Guillery. (2000). Economic Incentives for Rainforest Conservation Across Scales. *Science*. 288(5472):1828-1832.
- Langholz, J., Lassoie, J., & Schelhas, J. 2000. Incentives for Biological Conservation: Costa Rica's Private Wildlife Refuge, 14(6), 1735-1743.
- Landell-Mills, N., & Porras, I.T. 2002. Silver bullet or fools' gold? A global review of markets for forests environmental services and their impact on the poor. London: International Institute for Environment and Development.
- Lawry, S., Samii, C., Hall, R., Leopold, A., Hornby, D., & Mtero, F. (2013). The impact of land property rights interventions on investment and agricultural productivity in developing areas: a systematic review. Oslo: The Campbell Collaboration.
- Lawry, S. (2012). Devolution of Forest Rights and Sustainable Forest Management, Volumes I and II. Washington, DC: USAID.
- Lipscomb, M., & Mobarak, A. M. (2011). Decentralization and the Political Economy of Water Pollution: Evidence from the Re-drawing of County Border in Brazil. Unpublished manuscript, University of Notre Dame and Yale University.
- *Liu, C., Lu, J., Yin, R. (2010). An Estimation of the Effects of China's Priority Forestry Programs on Farmers' Income. *Environmental Management*, 45, 526-540.
- Mapedza, E. (2006). Compromised Co-Management, Compromised Outcomes: Experiences from a Zimbabwean Forest. *Africa Development*, 31(2), 123-146.
- #McAfee, K., & Shapiro, E.N. (2010). Payments for Ecosystem Services in Mexico: Nature, Neoliberalism, Social Movements, and the State. *Annals of the Association of American Geographers*, 100(3), 1-21.
- #Muñoz-Piña, C., Guevara, A., Torres, J.M., & Braña, J. (2008). Paying for the hydrological services of Mexico's forests: Analysis, negotiations and results. *Ecological Economics*, 65, 725-736.
- Muradian, R., Arsel, M. Pellegrini, L., Adaman, F., Aguilar, B., Agarwal, B., Corbera, E., Ezzine de Blas, D., Farley, J., Froger, G., Garcia-Frapolli, E., Gomez-Baggethun, E., Gowdy, J., Kosoy, N., Le Coq, J.F., Leroy, P., May, P. Meral, P., Mibielli, P., Norgaard, R., Ozkaynak, B., Pascual, U., Pengue, W., Perez, M., Pesche, D., Pirard, R., Ramos-Martin, J., Rival, L., Saenz, F., Van Hecken, G., Vatn, A., Vira, B., & Urama, K. (2013). Payments for

Ecosystem Services and the Fatal Attraction of Win-Win Solutions. *Conservation Letters*, 6(4):274-279.

*Ogada, M.J. (2013). Forest Management Decentralization in Kenya: Effects on Household Farm Forestry Decisions in Kakamega. In Holden, S.T., Otsuka, K., & Deininger, K., eds. *Land Tenure Reforms in Asia and Africa: Assessing Impacts on Poverty and Natural Resource Management*, 407-441. London: Palgrave-MacMillan.

#Ongugo, P.O., Mogoi, J.N., Obonyo, E., & Oeba, V.O. (2008). Examining the Roles of Community Forest Associations (CFAs) in the Decentralization Process of Kenyan Forests. Presented at the 12th Biennial Conference of the International Association for the Study of Commons (IASC), England.

Ostrom, E. (1990). *Governing the Commons: The Evolution of Institutions for Collective Action*. Cambridge: Cambridge University Press.

Ostrom, E. (2007). A Diagnostic Approach for Going Beyond Panaceas. *Proceedings of the National Academy of Sciences*, 104(39), 15181-15187.

#Pagiola, S. (2008). Payments for environmental services in Costa Rica. *Ecological Economics*, 65, 712-724.

#Pagiola, S., Arcenas, A., & Platais, G. (2004). Can Payments for Environmental Services Help Reduce Poverty? An Exploration of the Issues and Evidence to Data from Latin America. *World Development* 33(2), 237-253.

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(2), 254-274.

Porrás, I. (2010). *Fair and Green? Social Impacts of Payments for Environmental Services in Costa Rica*. London: International Institute for Environment and Development.

Potter, L., & Badcock, S. (2001). The Effects of Indonesia's Decentralisation of Forests and Estate Crops in Riau Province: Case Studies of the Original Districts of Kampar and Indragiri Hulu. Bogor: CIFOR.

Pullin, A., Bangpan, M., Dalrymple, S., Dickson, K., Healey, J., Hockley, N., Jones, J., Knight, T., & Oliver, S. (2013). Systematic Review: Human Well-being Impacts of Terrestrial Protected Areas. *Environmental Evidence*, 2(19) (online at <http://www.environmentalevidencejournal.org/content/2/1/19>).

Puryavaud, J.P. 2003. Standardizing the calculation of the annual rate of deforestation. *Forest Ecology and Management*, 177, 593-596.

Ravallion, M., Datt, G., & van de Walle, D. (1991). Quantifying Absolute Poverty in the Developing World. *Review of Income and Wealth*, 37(4), 345-361.

*Robalino, J., & Pfaff, A. (2013). Ecopayments and Deforestation in Costa Rica: A Nationwide Analysis of PSA's Initial Years. *Land Economics*, 89(3):432-448.

*Robalino, J., Pfaff, A., Sánchez-Azofeifa, G.A., Alpízar, F., León, C., Rodríguez, C.M. (2008). Deforestation Impacts of Environmental Services Payments: Costa Rica's PSA Program 2000-2005. Discussion Paper Series, Environment for Development.

- Robalino, J., & Villalobos-Fiatt, L. (2010). Conservation policies and labor markers: unraveling the effects of national parks on local wages in Costa Rica. *Environment for Development Discussion Paper Series*, 10-02.
- Rojas, M. & Aylward, B. 2003. What are we learning from experiences with markets for environmental services in Costa Rica? Hertfordshire: International Institute for Environment and Development.
- Salzman, J. (2010). Designing Payments for Ecosystem Services. Working Paper, PERC Policy Series, No. 48.
- Samii, C., Lisiecki, M., Kulkarni, P., Paler, L., & Chavis, L. (2013). Protocol: Effects of Payment for Environmental Services and Decentralized Forest Management on Deforestation and Poverty in Low and Middle Income Countries: A Systematic Review. London: The Campbell Collaboration International Development Coordinating Group.
- Samuelson, P. (1954). The Pure Theory of Public Expenditure. *Review of Economics and Statistics*, 36(4), 387-389.
- *Sánchez-Azofeifa, G.A., Pfaff, A., Robalino, J.A., & Boomhower, J.P. (2007). Costa Rica's Payment for Environmental Services Program: Intention, Implementation, and Impact. *Conservation Biology*, 21(5), 1165-1173.
- Santilli, M., Moutinho, P., Schwartzman, S., Nepstad, D., Curran, L., & Nobre, C. (2005). Tropical Deforestation and the Kyoto Protocol. *Climatic Change*, 71, 267-276.
- *Scullion, J., Thomas, C.W., Vogt, K.A., Pérez-Maqueo, O., & Logsdon, M.G. (2011). Evaluating the environmental impact of payments for ecosystem services in Coatepec (Mexico) using remote sensing and on-site interviews. *Environmental Conservation*, 38(4), 426-434.
- Sims, K.R.E. (2010) Conservation and development: Evidence from Thai protected areas. *Journal of Environmental Economics and Management*, 60, 94-114.
- *Somanathan, E., Prabhakar, R., & Mehta, B.S. (2009). Decentralization for Cost-Effective Conservation. *Proceedings of the National Academy of Sciences*, 106(11):4143-4147.
- Sunderlin, W., Angelsen, A., Belcher, B., Burgers, P., Nasi, R., Santoso, L., & Wunder, S. (2005). Livelihoods, forests, and conservation in developing countries. *World Development*, 33(9), 1383-1402.
- Tacconi, L. (2007). Decentralization, Forests and Livelihoods: Theory and Narrative. *Global Environmental Change*, 17, 338-348.
- Vajja, A. & White, H. (2008). Can the World Bank Build Social Capital? The Experience of Social Funds in Malawi and Zambia. *Journal of Development Studies*, 44(8), 1145-1168.
- Vandermeer, J., & Perfecto, I. (1995). *Breakfast of Biodiversity: The Truth About Rain Forest Destruction*. Oakland, CA: Food First.
- Van Kooten, C., & Sohngen, B. (2007). Economics of Forest Ecosystem Carbon Sinks: A Review. *International Review of Environmental and Resource Economics*, 1(3), 237-269.

- Viechtbauer, W. (2010). Conducting Meta-Analyses in R with the metafor Package. *Journal of Statistical Software*, 36(3), (online access).
- West, P.W. (2009). *Tree and Forest Measurement*, Second Edition. Heidelberg: Springer-Verlag.
- Wilson, D.B., Weisburd, D., & McClure, D. (2011). Use of DNA testing in police investigative work for increasing offender identification, arrest, conviction and case clearance. *Campbell Systematic Reviews* 2011(7).
- Wu, J., Zilberman, D., & Babcock, B. (2001). Environmental and distributional impacts of conservation targeting strategies. *Journal of Environmental Economics and Management* 41, 333-350.
- Wunder, S. (2001). Poverty alleviation and tropical forests: What scope for synergies? *World Development*, 29(11), 1817-1833.
- Wunder, S. (2005). Payments for environmental services: Some nuts and bolts. *CIFOR Occasional Paper*, 42.
- Wunder, S. (2007). The efficiency of payment for environmental services in tropical conservation. *Conservation Biology*, 21(1), 8-58.
- Wunder, S. (2013). When payments for environmental services will work for conservation. *Conservation Letters*, 6(4):230-237.
- #Wunder, S., Engel, S., & Pagiola, S. (2008). Taking stock: A comparative analysis of payments for environmental services programs in developed and developing countries. *Ecological Economics*, 65, 834-852.
- #Zbinden, S., & Lee, D.R. (2005). Paying for Environmental Services: An Analysis Of Participation in Costa Rica's PSA Program. *World Development*, 33(2), 255-272

Studies included for quantitative synthesis

PES

- Alix-Garcia, J., Shapiro, E.N., & Sims, K.R.E. (2012). Forest Conservation and Slippage: Evidence from Mexico's National Payments for Ecosystem Services Program. *Land Economics*, 88(4), 613-638.
- Arriagada, R.A., Sills, E.O., & Pattanayak, S.K. (2011). Payments for Environmental Services and Their Impact on Forest Transition in Costa Rica. Working Paper.
- Arriagada, R.A., Ferraro, P.J., Sills, E.O., & Pattanayak, S.K. (2012). Do Payments for Environmental Services Affect Forest Cover?: A Farm-Level Evaluation from Costa Rica. *Land Economics*, 88(2), 382-399.
- Garbach, K., Lubell, M., & DeClerck, F.A.J. (2012). Payment for Ecosystem Services: The roles of positive incentives and information sharing in stimulating adoption of silvopastoral conservation practices. *Agriculture, Ecosystems and Environment*, 156, 27-36.

Hegde, R., & Bull, G.Q. (2011). Performance of an agro-forestry based Payments-for-Environmental-Services project in Mozambique: A household level analysis. *Ecological Economics*, 71, 122-130.

Honey-Rosés, J., Baylis, K., & Ramírez, M.I. (2011). A Spatially Explicit Estimate of Avoided Forest Loss. *Conservation Biology*, 25(5), 1032-1043.

Liu, C., Lu, J., Yin, R. (2010). An Estimation of the Effects of China's Priority Forestry Programs on Farmers' Income. *Environmental Management*, 45, 526-540.

Robalino, J., & Pfaff, A. (2013). Ecopayments and Deforestation in Costa Rica: A Nationwide Analysis of PSA's Initial Years. *Land Economics*, 89(3):432-448.

Robalino, J., Pfaff, A., Sánchez-Azofeifa, G.A., Alpízar, F., León, C., Rodríguez, C.M. (2008). Deforestation Impacts of Environmental Services Payments: Costa Rica's PSA Program 2000-2005. Discussion Paper Series, Environment for Development.

Sánchez-Azofeifa, G.A., Pfaff, A., Robalino, J.A., & Boomhower, J.P. (2007). Costa Rica's Payment for Environmental Services Program: Intention, Implementation, and Impact. *Conservation Biology*, 21(5), 1165-1173.

Scullion, J., Thomas, C.W., Vogt, K.A., Pérez-Maqueo, O., & Logsdon, M.G. (2011). Evaluating the environmental impact of payments for ecosystem services in Coatepec (Mexico) using remote sensing and on-site interviews. *Environmental Conservation*, 38(4), 426-434.

ADDITIONAL STUDIES included for DATA ON CONTEXT AND INSTITUTIONS

Arriagada, R.A., Sills, E.O., Pattanayak, S.K., & Ferraro, P.J. (2009). Combining Qualitative and Quantitative Methods to Evaluate Participation in Costa Rica's Program of Payments for Environmental Services. *Journal of Sustainable Forestry*, 28, 343-367.

Cole, R.J. (2010). Social and environmental impacts of payments for environmental services for agroforestry on small-scale farms in southern Costa Rica. *International Journal of Sustainable Development & World Ecology*, 17(3), 208-216.

Honey-Rosés, J., López-García, J. Rendón-Salinas, E., Peralta-Higuera, A., & Galindo-Leal, C. (2009). To pay or not to pay? Monitoring performance and enforcing conditionality when paying for forest conservation in Mexico. *Environmental Conservation*, 36(2), 120-128.

McAfee, K., & Shapiro, E.N. (2010). Payments for Ecosystem Services in Mexico: Nature, Neoliberalism, Social Movements, and the State. *Annals of the Association of American Geographers*, 100(3), 1-21.

Muñoz-Piña, C., Guevara, A., Torres, J.M., & Braña, J. (2008). Paying for the hydrological services of Mexico's forests: Analysis, negotiations and results. *Ecological Economics*, 65, 725-736.

Pagiola, S. (2008). Payments for environmental services in Costa Rica. *Ecological Economics*, 65, 712-724.

Pagiola, S., Arcenas, A., & Platais, G. (2005). Can Payments for Environmental Services Help Reduce Poverty? An Exploration of the Issues and the Evidence to Date from Latin America. *World Development*, 33(2), 237-253.

Wunder, S., Engel, S., & Pagiola, S. (2008). Taking stock: A comparative analysis of payments for environmental services programs in developed and developing countries. *Ecological Economics*, 65, 834-852.

Zbinden, S., & Lee, D.R. (2005). Paying for Environmental Services: An Analysis Of Participation in Costa Rica's PSA Program. *World Development*, 33(2), 255-272

Appendices

Search Strategy Appendix

Electronic search databases and websites

We reproduce the list of sources from Bowler et al. (2010, pp. 55-56):

Literature databases

- AgEcon (<http://ageconsearch.umn.edu/>)
- Greenfile (Ebsco)
- Agris (FAO - <http://agris.fao.org/>)
- RUPES (<http://rupes.worldagroforestry.org>)
- Science and Social Science Citation Index
- British Library for Development Studies
- Scopus
- Agricola
- CAB Abstracts
- EMBASE
- Science Direct
- EconLit
- JSTOR
- Directory of Open Access Journals
- IDEAS

Web search engines [NB: “jux2.com” is excluded from the original list]:

- <http://www.google.com>
- <http://scholar.google.com>
- <http://scientific.thomsonwebplus.com/>
- <http://www.scirus.com> (restricted to “web sources” only)

Specialist websites

- <http://www.capri.cgiar.org/>
- <http://www.catie.org.ac.cr/>
- <http://www.cbnrm.net/>
- <http://www.cgiar.org/>

- <http://www.cifor.cgiar.org>
- <http://www.cof.orst.edu/org/istf/ftpp.htm>
- <http://www.communityforestryinternational.org/>
- <http://www.conservation.org>
- <http://www.dfid.gov.uk>
- <http://www.etfrn.org>
- <http://www.forestrycenter.org/>
- <http://forests.org/>
- <http://www.forestsandcommunities.org/>
- <http://www.ifad.org/>
- <http://www.iied.org>
- <http://www.indiana.edu/~iascp/>
- <http://www.iucn.org>
- <http://www.livelihoods.org>
- <http://www.www.macp-pk.org>
- <http://www.odi.org>
- <http://www.www.panda.org>
- <http://www.pfc.cfs.nrcan.gc.ca/>
- <http://www.rainforestportal.org/>
- <http://www.recoftc.org>
- <http://www.thegef.org>
- <http://www.tropenbos.nl/>
- <http://www.usaid.gov/>
- <http://www.waldbau.uni-freiburg.de/forlive/Home.html>
- <http://www.wcs.org>

Specialist agencies contacted via email

- United Nations Development Programme (UNDP)
- United Nations Environment Programme (UNEP)
- World Bank
- African Development Bank (AFDB)

- Asian Development Bank (ADB)
- European Bank for Reconstruction and Development (EBRD)
- Inter-American Development Bank (IDB)
- International Fund for Agricultural Development (IFAD)
- UN Food and Agriculture Organisation (FAO)
- UN Industrial Development Organisation (UNIDO)

Field journals

American Economic Review

American Economic Journal: Applied

American Economic Journal: Economic Policy

American Journal of Political Science

American Political Science Review

Conservation Biology

Ecological Economics

Environment and Development Economics

Environment, Development and Sustainability

Environmental and Resource Economics

Forest Policy and Economics

Journal of Development Economics

Journal of Environmental Economics and Management

Journal of Environmental Planning and Management

Journal of Forest Economics

Journal of Politics

Journal of Public Economics

Journal of Regulatory Economics

Journal of Sustainable Forestry

Land Economics

Proceedings of the National Academy of Sciences

Resource and Energy Economics

Review of Economics and Statistics

Review of Environmental Economics and Policy

World Development

LMIC filter

Below is the set of terms used to filter searches and limit results to studies carried out in low or middle income countries (LMICs):

AND ("Africa" OR "Asia" OR "Caribbean" OR "West Indies" OR "South America" OR "Latin America" OR "Central America" OR "Afghanistan" OR "Albania" OR "Algeria" OR "American Samoa" OR "Angola" OR "Argentina" OR "Armenia" OR "Azerbaijan" OR "Bangladesh" OR "Benin" OR "Belize" OR "Bhutan" OR "Bolivia" OR "Botswana" OR "Brazil" OR "Bulgaria" OR "Burkina Faso" OR "Burundi" OR "Cambodia" OR "Cameroon" OR "Cape Verde" OR "Central African Republic" OR "Chad" OR "Chile" OR "China" OR "Colombia" OR "Comoros" OR "Congo" OR "Costa Rica" OR "Cote d'Ivoire" OR "Cuba" OR "Djibouti" OR "Dominica" OR "Dominican Republic" OR "East Timor" OR "Ecuador" OR "Egypt" OR "El Salvador" OR "Eritrea" OR "Ethiopia" OR "Fiji" OR "Gabon" OR "Gambia" OR "Ghana" OR "Grenada" OR "Guatemala" OR "Guinea" OR "Guinea-Bissau" OR "Guam" OR "Guyana" OR "Haiti" OR "Honduras" OR "India" OR "Indonesia" OR "Ivory Coast" OR "Jamaica" OR "Jordan" OR "Kazakhstan" OR "Kenya" OR "Kyrgyzstan" OR "Laos" OR "Lebanon" OR "Lesotho" OR "Liberia" OR "Madagascar" OR "Malaysia" OR "Malawi" OR "Mali" OR "Malta" OR "Mauritania" OR "Mauritius" OR "Mexico" OR "Micronesia" OR "Moldova" OR "Mongolia" OR "Morocco" OR "Mozambique" OR "Myanmar" OR "Namibia" OR "Nepal" OR "Nicaragua" OR "Niger" OR "Nigeria" OR "Pakistan" OR "Panama" OR "Papua New Guinea" OR "Paraguay" OR "Peru" OR "Philippines" OR "Puerto Rico" OR "Rwanda" OR "Senegal" OR "Sierra Leone" OR "Sri Lanka" OR "Somalia" OR "Sudan" OR "Swaziland" OR "Tajikistan" OR "Tanzania" OR "Thailand" OR "Togo" OR "Tonga" OR "Tunisia" OR "Turkey" OR "Turkmenistan" OR "Uganda" OR "Uzbekistan" OR "Venezuela" OR "Vietnam" OR "Yemen" OR "Zambia" OR "Zimbabwe")

Example of a detailed search strategy

To illustrate how the search strategy was applied, here are the steps used to search the AgEcon database (<http://ageconsearch.umn.edu/>) as applied on February 16, 2013:

- Go to database website: <http://ageconsearch.umn.edu/>
- Enter into the search text boxes the following:
 - Box 1: "pay* OR reward* OR incentiv* OR compensat*" with search type "anywhere in record"
 - Box 2: AND "forest OR deforest* OR ecol* OR ecos* environment* OR conservation" with search type anywhere in record.
 - Box 3: AND enter sections of the LMIC filter shown above (the entire filter cannot be entered at once, so enter sections of the filter until all keywords have been used) with search type "anywhere in record"
- The search yields 185 hits with title and abstract information. Extract information and enter in search database (using Endnote).

Brief descriptions of programs included in quantitative synthesis

PES

Alix-Garcia et al. 2012: In 2003, Mexico began paying landowners to maintain forest cover through their PSAH (Pago por Servicios Ambientales-Hidrológico) program. This program signs participants to 5-year contracts under which they receive payments (on a per hectare basis) for maintaining forest cover on enrolled land.

Arriagada et al. 2011: The Costa Rican Forestry Law 7575 of 1996 established the PSA (Pagos por Servicios Ambientales) program. Under PSA, landowners sign contracts for either forest conservation (no land-cover change allowed over a 5-year period), reforestation (plant trees on agricultural/abandoned land and maintain for 15 years), or sustainable forest management (prepare 'sustainable logging plan' for low-intensity logging while keeping forests intact and follow it for 15 years).

Arriagada et al. 2012: The Costa Rican Forestry Law 7575 of 1996 established the PSA (Pagos por Servicios Ambientales) program. Under PSA, landowners sign contracts for forest conservation (no land-cover change allowed over a five year period), reforestation (plant trees on agricultural/abandoned land and maintain for 15 years), or sustainable forest management (prepare 'sustainable logging plan' for low-intensity logging while keeping forests intact and follow it for 15 years).

Garbach et al. 2012: In 2002, Costa Rica established the Regional Integrated Silvopastoral Approaches to Ecosystem Management Project (RISEMP). Under RISEMP, landowners received payments for adopting any of the following sustainable land management practices: improving pasture by introducing hybrid grass species, planting a live fence, planting trees in existing pastureland (either low- or high-density, planting shrubs adjacent to pasture/crop fields, and planting forests around water sources.

Hegde & Bull 2011: In Mozambique, a PES program is underway in the Nhambita Community (located in Chicale Regulado). Under this program, farmers sign contracts with NGOs to plant trees on their farm, which they manage for 25 years in exchange for conditional cash payments.

Honey-Rosés et al. 2011: In order to protect the winter habitat of monarch butterflies, in 2000, Mexico established the Monarch Butterfly Conservation Fund (MBCF). This program pays landowners in the monarch butterfly habitat to abstain from logging.

Liu et al. 2010: In 1999, China established the Sloping Land Conversion Program (SLCP). Under this program, the central government provides farmers with seeds/seedlings, grain, and cash in exchange for converting sloping or desertified cropland into ecological/economic forest or grassland.

Robalino and Pfaff 2013: The Costa Rican Forestry Law 7575 of 1996 established the PSA (Pagos por Servicios Ambientales) program. Under PSA, landowners sign contracts for either forest conservation (no land-cover change allowed over a 5-year period), reforestation (plant trees on agricultural/abandoned land and maintain for 15 years), or sustainable forest management (prepare "sustainable logging plan" for low-intensity logging while keeping forests intact and follow it for 15 years).

Robalino et al. 2008: The Costa Rican Forestry Law 7575 of 1996 established the PSA (Pagos por Servicios Ambientales) program. Under PSA, landowners sign contracts for forest conservation (no land-cover change allowed over a 5-year period), reforestation (plant trees on agricultural/abandoned land and maintain for 15 years), or sustainable forest management (prepare “sustainable logging plan” for low-intensity logging while keeping forests intact and follow it for 15 years).

Sánchez-Azofeifa et al. 2007: The Costa Rican Forestry Law 7575 of 1996 established the PSA (Pagos por Servicios Ambientales) program. Under PSA, landowners sign contracts for forest conservation (no land-cover change allowed over a 5-year period), reforestation (plant trees on agricultural/abandoned land and maintain for 15 years), or sustainable forest management (prepare ‘sustainable logging plan’ for low-intensity logging while keeping forests intact and follow it for 15 years).

Scullion et al. 2011: In 2003, Mexico began paying landowners to maintain forest cover through their PSAH (Pago por Servicios Ambientales-Hidrológico) program. This program signs participants to 5-year contracts under which they receive payments (on a per hectare basis) for maintaining forest cover on enrolled land.

Effect sizes

Imputing standard errors

In some cases, standard errors are not reported but rather t-statistics, p-values, or sometimes only significance levels. When t-statistics were reported for an effect Δ , we computed the standard error as Δ / t . From significance levels, we imputed the standard error from a t-statistic equal to the quantile at the posted significance level---e.g., if an effect Δ was shown to have $p < .05$ for a two-way test, we imputed a t-statistic corresponding to the .975 quantile of the normal distribution ($t = 1.96$) and then a standard error corresponding to $|\Delta / t|$. Generally speaking the formula for imputed standard errors ($se.imp$) from a two-sided p value under a normal approximation is as follows:

$$se.imp = \Delta / \Phi^{-1}(1-.5*p),$$

where Φ^{-1} is the inverse CDF of the normal distribution

When no standard error, t-statistic, or statistical significance level was given, we imputed a p-value of 0.5 and then assigned the associated standard error, which is equivalent to assigned a standard error equal to $(1/0.67) |\Delta| = 1.48 |\Delta|$. Imputing a p-value of 0.5 is not completely arbitrary, as it corresponds to the mean of the posterior p-value distribution under the null hypothesis, given a uniform prior over 0 to 1. In addition, such constant scaling will mechanically impute smaller standard errors for estimates closer to zero, in which case inverse weighted averages across numerous estimates will tend to drive the average toward zero; this again is consistent with assuming a prior of a null effect and updating it with vague information.

Standardized forest cover effect sizes

Puyravaud (2003) proposes a standardized measure of forest cover change based on the compound interest law,

$$C = C_0 \exp[r(t_2-t_1)]$$

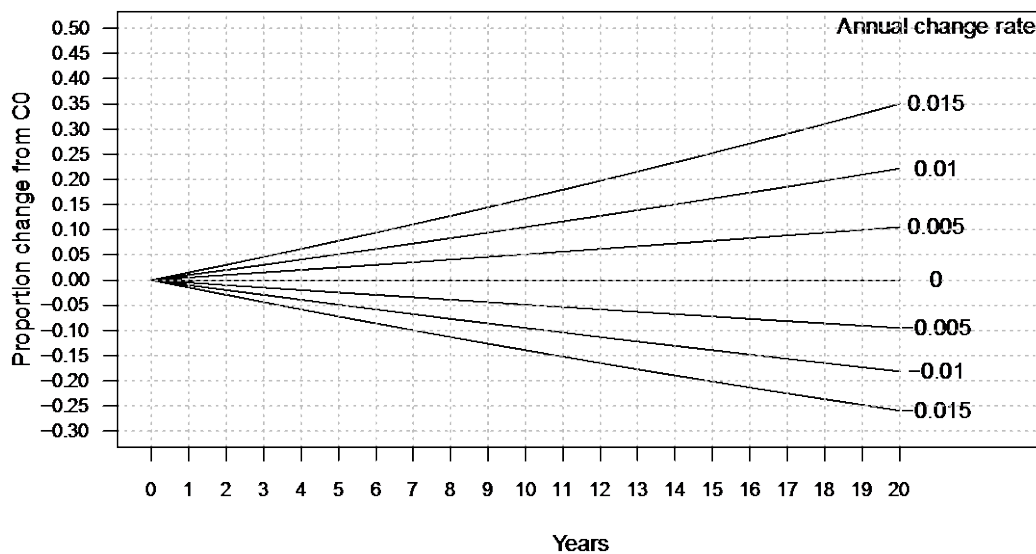
where C is the amount of forest cover at the time of follow-up, C₀ is forest cover at baseline, r is the continuous rate of change per unit of time, and t₂-t₁ is the amount of time elapsed between periods t₁ and t₂. Taking the natural log of both sides and rearranging yields

$$r = \ln(C/C_0)/(t_2-t_1)$$

This measure of rate of change takes a sign that is positive for net forest cover growth and negative for net deforestation. The quantity 100r% is interpretable as the percent change in forest cover per period. For the studies considered above, we use year as the relevant period. Figure A.1 below shows how this annual rate of change translates into proportion change in forest cover for up to twenty years. Thus, a program that has the effect of sustaining a .01 increase in the annual rate of forest cover change (or, a .01 decrease in the deforestation rate) would induce on

the order of a 10 percent increase in the extent of forest cover after ten years and 20 percent increase in the extent of forest cover after twenty years, as compared to the counterfactual of no program (at these small values of r, the annual change rate, and for these time scales, the compound interest law is almost perfectly linear in time).

Figure A. 1: Forest cover change under the compound interest law.



Moving from this measure of forest cover change to a standardized effect measure may proceed as follows. We work with the difference between the actual forest cover change rate in the treated area and the counterfactual change rate for that area. Studies typically report forest coverage on an average-per-parcel basis. Given N parcels, then this does not affect the calculations as $(C/N)/(C_0/N) = C/C_0$. Using the a subscript to denote quantities for the actual treated area and c subscript for counterfactual quantities, then we note that

$$r_a - r_c = [\ln(C_a/C_0) - \ln(C_c/C_0)]/(t_2-t_1)$$

$$= [\ln(C_a/C_c)]/(t_2-t_1)$$

$$= [\ln((C_a/N)/\{(C_a/N) - \Delta\})]/(t_2-t_1),$$

where Δ is the estimated effect on mean forest cover change (in area units) in the N parcels, C_a/N is mean forest cover in the N treated parcels and $(C_a/N) - \Delta$ estimates mean counterfactual forest cover in the treated parcels. Given a standard error for Δ denoted as $se(\Delta)$, an approximate standard error for the difference in rates that takes the treated area forest cover C_a as fixed is obtained via the delta method as

$$se(r_a - r_c) = se(\Delta)/\{(C_a/N) - \Delta\}(t_2 - t_1).$$

For studies that report effects in terms of proportion of fully forested parcels deforested, denoted as Δ_p , we have that the average pre-treatment forest cover in treated parcels,

$$P_0 = (C_0/N)/(A/N) = C_0/A,$$

is fixed to 1, where A denotes the sum of parcel areas. To compute $r_a - r_c$, we need the average actual post-treatment forest cover proportion in treated areas,

$$P_a = (C_a/N)/(A/N),$$

which we can also use to compute the counterfactual forest cover proportion,

$$P_c = (C_c/N)/(A/N) = P_a - \Delta_p.$$

Since $P_a/P_c = C_a/C_c$, we have

$$r_a - r_c = \{\ln[P_a/(P_c - \Delta_p)]\}/(t_2 - t_1),$$

with approximate standard error

$$se_p(r_a - r_c) = se(\Delta_p)/[(P_c - \Delta_p)(t_2 - t_1)].$$

In cases where P_a is not reported, we impute a value using the treatment parcels deforestation rate in the most similar case where such information is provided.

Standardized consumption and income effects

We standardized consumption or income effects in terms of percentage change relative to the counterfactual. For studies that estimate effects using log of income or log of consumption expenditure as the outcome, then for an effect estimated as Δ_l , the percentage change over the counterfactual is given by

$$PC = 100[\exp(\Delta_l) - 1]\%.$$

A delta method approximate standard error is given by

$$se_e(PC) = 100 se(\Delta_l) \exp(\Delta_l).$$

For studies that use the raw income or consumption expenditure levels as the outcome, then for an effect estimated as Δ_r , the percentage change over the counterfactual is given by

$$PC = 100[\Delta_r / (T - \Delta_r)]\%,$$

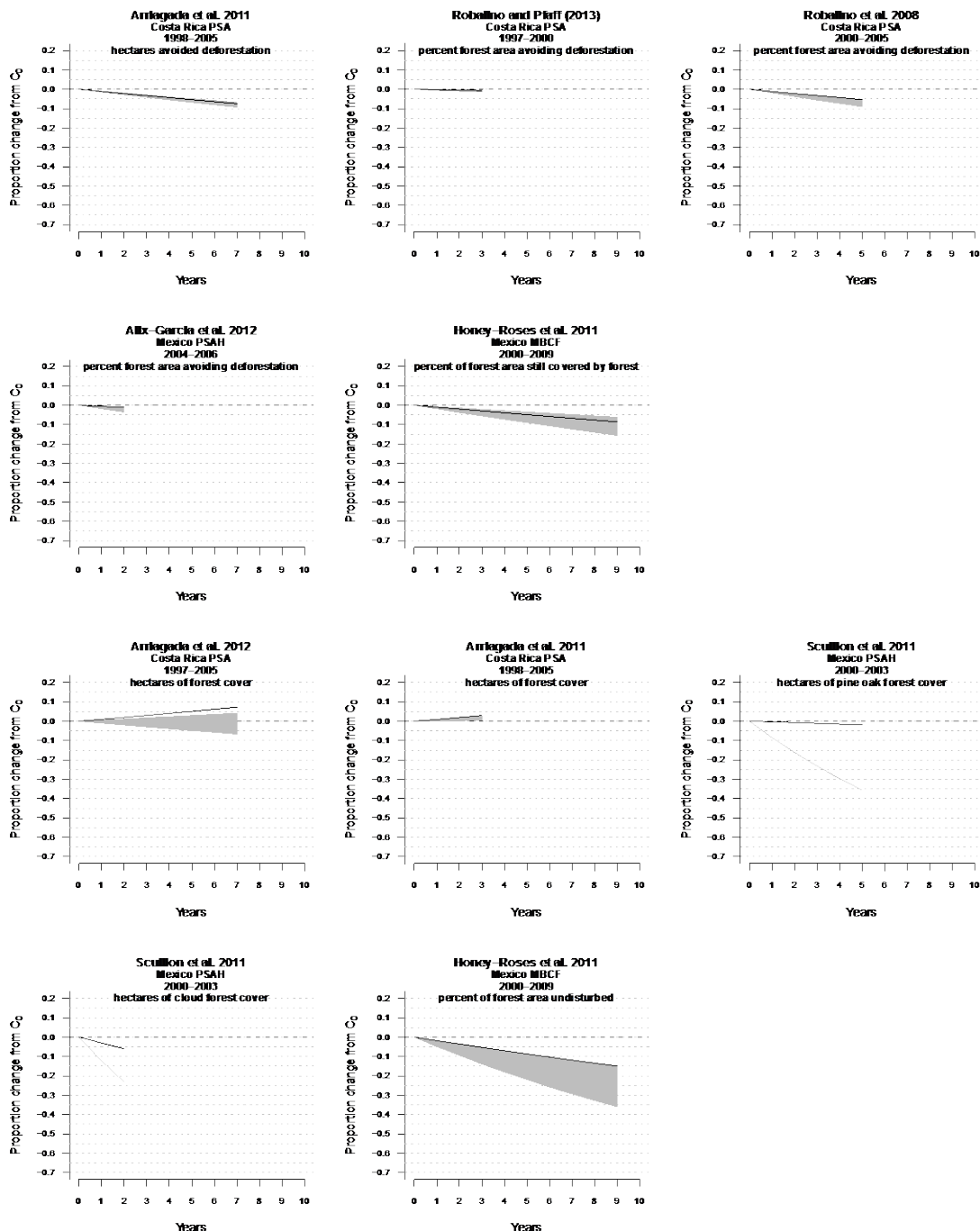
where T is the mean income level in the treatment group. A delta method approximate standard error is given by,

$$se_p(PC) = 100 se(\Delta_r) [(T) / (T - \Delta_r)^2].$$

Implications of Rate of Change Effects for Forest Cover Trajectories

Figure A.2 and A.3 show the implications of the rate-of-change effects for forest cover trajectories. The x axis shows years. The y axis shows proportional change in forest cover relative to the amount of forest cover that prevailed before the program was implemented (this baseline level of forest cover is denoted as C_0 , refer to the discussion in appendix section 9.3.2 on standardized forest cover change measures). A horizontal dashed reference line is drawn at 0 on the y axis. This reference line would correspond to no forest cover change over time. The black curves trace out the actual forest cover change trajectories in the program areas (treatment group) as reported in each of the studies. We trace out the change trajectory for the number of years that the program ran before the assessment provided by the study. Trajectories that run below the zero reference line imply forest loss; trends that run above the zero reference line imply forest gain. Each graph also displays a gray shaded area that corresponds to the 95% confidence interval for the estimated counterfactual change trajectory. That is, the gray shaded area translates the effect estimate from the study ($r_a - r_c$, in the notation from appendix section 9.3.2) into an estimate of what would have happened in the program areas had there been no program (thus, the counterfactual). If the black trajectory line overlaps with the gray shaded area, this means that the study found no statistically significant effect (at 95% confidence). When the black line does not overlap with the gray shaded area, this means that the effect of the program was statistically significant and thus the implied counterfactual trajectory is clearly distinct from what actually transpired in the program area. The titles for each graph show the study authors, the program, the observation period from which the estimates were derived, and the outcome that was used in the original analysis.

Figure A. 2: Implications of estimated effects of PES for forest cover change. The y axis measures proportional change in forest cover relative to the amount of forest cover prior to the intervention (with this baseline forest cover denoted as C_0). The black lines trace out forest cover trends in the program (“treated”) areas, and gray shaded areas show the 95% confidence interval for the counterfactual forest cover trends implied by the effect estimates shown in Table 5 and Figure 4. The first five graphs show effects on forest cover attributable to deforestation (in which case, trajectories can never go above zero), the next five on forest attributable to either deforestation or forest growth. Effect estimates from Scullion et al. (2011) were not accompanied by standard errors and so we simply trace out the point estimate for the counterfactual forest cover trend.



Deviations from protocol

Our protocol proposed that our risk of bias assessment code studies as 'high', 'low', or 'unclear' risk of bias for each of the domains considered (exogenous assignment, control for confounding, avoidance of motivation bias, accounting for spill-over, avoidance of selective outcome reporting, avoidance of analysis fishing, and appropriate statistical inference). We decided rather to indicate as to whether the study satisfied these criteria by indicating 'yes', 'no', or 'unclear', which can be interpreted as equivalent to the designations of 'low', 'high', and 'unclear' risk of bias with respect to each of these domains.

Our protocol specified meta-regression analyses to test our moderator and mediator hypotheses and a set of quantitative publication bias analyses. We were unable to implement these as the number of eligible quantitative studies was too few.

The protocol also included a proposal for a set of descriptive and moderation analyses to assess external validity of our estimates. We were unable to implement these as the number of eligible studies and associated contexts was too few.

The study was initially designed to also include Decentralised Forest Management (DFM) programs, as outlined in the ToR from the funder and Samii et al (2013). However, the search, data extraction and analysis was conducted in parallel rather than integrated and once the review was completed it was decided to split the two interventions into two separate reviews for ease of interpretation.

Study coding form

ITEMS/QUESTIONS LISTED IN THIS COLUMN	ENTER RESPONSES IN THIS COLUMN	INSTRUCTIONS
Coder Information		
Name of person filling in this form		First Last
Date that the form was begun		MM/DD/YY
General Study Information		
Author 1		First Last
Author 2 (if applicable)		First Last
Author 3 (if applicable) (If more than 3 authors, include only the first 3.)		First Last
Year of publication		YYYY
Intervention type		▲ PES or DFM
Publication type		▲ Refer to the final-most publication. This is important because it determines whether the study was subject to peer review. Journal articles and book chapters should be straightforward to identify. Working papers will bear the mark of a working paper series. Reports will bear the mark of a commissioning or sponsoring organization. Unpublished works bearing no such series or organizational marks are considered as unpublished.
Study Context		
Country		Country name
Year the intervention started		YYYY
Year of final outcome assessment/final data collection		YYYY
General Eligibility		
Does the paper report results of primary research on the impact of a PES or DFM program?		▲ Primary research means that it contains original data analysis, as opposed to summarizing analyses by others.
Does the study assess impacts on at least one of the following: a. deforestation, b. material welfare in terms of poverty, consumption, or income?		▲ These are the outcomes of interest for our review.
Does the study examine a developing country?		▲ Cf. the World Bank list of Low Income, Lower Middle Income, or Upper Middle Income countries: http://data.worldbank.org/about/country-classifications/country-and-lending-groups
5. Does the study provide information on all of the following? a. research questions; b. data collection procedures; c. demographic or contextual characteristics of the subjects studied; d. analysis methods.		▲ Cf. the primary write-up of the study as well as any supporting materials (e.g., appendices, accompanying reports, etc.)

PROCEED ONLY IF YOU ANSWERED "YES" TO ALL OF THE ABOVE.		
Eligibility for Quantitative Synthesis		
Were any of the following experimental or quasi-experimental methodologies used to assess impact of PES or DFM on deforestation or welfare (poverty, income, or consumption)? (Mark yes for each that applies.)		<i>These methods should be explicitly mentioned in the study write-up. If not, and if it is unclear whether any of the methods below were used, indicate "unclear", and then we can try to contact the authors.</i>
...randomized experiment or randomized control trial	<input type="checkbox"/>	▲▼
...regression adjustment to control for confounding variables	<input type="checkbox"/>	▲▼
...difference-in-differences methodology	<input type="checkbox"/>	▲▼
...instrumental variables methodology	<input type="checkbox"/>	▲▼
...panel/fixed-effects methodology	<input type="checkbox"/>	▲▼
...regression discontinuity methodology	<input type="checkbox"/>	▲▼
...matching or weighting methodology	<input type="checkbox"/>	▲▼
Were quantitative impact estimates reported on the following? (Mark yes for each that applies.)		<i>By "reported" we mean that they appear clearly in some kind of table or graph.</i>
...deforestation	<input type="checkbox"/>	▲▼
...income, consumption or poverty	<input type="checkbox"/>	▲▼
Is the study eligible for quantitative synthesis? (Mark yes <i>only</i> if an experimental or quasi-experimental methodology was used AND impact estimates were reported on at least one of the outcomes above.)	<input type="checkbox"/>	▲▼ <i>The study is ineligible if you answered "no" to all of the above.</i>
Eligibility for Qualitative Synthesis		
Is the aim of the study clearly about the impact of PES or DFM?	<input type="checkbox"/>	▲▼
Does the study work from a theoretical framework?	<input type="checkbox"/>	▲▼ <i>This would be to distinguish a qualitative study from, say, a journalistic account or a report that is mostly intended to advertise rather than scrutinise.</i>
Are original qualitative or quantitative data (e.g., quotes/paraphrasing from interviews, close observation, process tracing, etc.) used to support conclusions about impact, background conditions, or mediating factors?	<input type="checkbox"/>	▲▼
Is the study eligible for qualitative synthesis? (Mark yes <i>only</i> if you answered yes to all three of the questions above.)	<input type="checkbox"/>	▲▼

PROCEED ONLY IF THE STUDY QUALIFIES FOR QUANTITATIVE OR QUALITATIVE SYNTHESIS		
Design of the Intervention		
At what level do we have measured variation in intervention assignment or levels of exposure to the intervention? (This is the "unit of intervention" for the sake of our analysis.)		That is, at what level do we have actual measured variation in program exposure or intensity. In some cases this is clear --- for example a household level program for which we have data from household about whether they indeed received the program. In other cases it is more subtle---for example, for programs that are the result of regional policies, the correct answer here is "region," even if they are targeted at households or communities within the region.
PROCEED ONLY IF THE STUDY QUALIFIES FOR QUANTITATIVE SYNTHESIS		
Design of Quantitative Study (skip if study is not eligible for quantitative synthesis)		
What were the units of observation/data collection?		That is, what are the "units" that are used in the data analysis. You can determine this by looking at the sample size. For example, if the intervention was applied in, say, 100 communities, but the data analysis uses data from 1000 households, then the unit of observation must be the household (while the unit of intervention would be the community).
Is this a clustered study? (Mark yes if the units of observation are nested within units of assignment---e.g., units of assignment are communities, but units of observation are households.)		Continuing with the example above, that would be a clustered study, with the "clusters" being the communities, and an average of 10 households per cluster (10 households per community*100 communities = 1000 households).
Does the analysis adequately account for any clustering in the design of the intervention or the study?		That is, does the study indicate that they used "cluster robust", "multilevel", or "hierarchical" statistical methods?
How many units of intervention were included in the study?		In our running example, this would be the 100 communities.
How many units of observation are included in the study?		In our running example, this would be the 1000 households.
How many different treatment/control conditions were studied?		This number should always be 2 or higher. In some cases, a single study will examine different combinations of interventions. If there is only one intervention type and it is compared to a "control" condition, then you want to write "2" here.
Describe the treatment conditions.		Provide a concise description of the way the treatments were defined for the purposes of the statistical analysis that was performed.
Were any significant imbalances in the baseline characteristics of the intervention recipient and non-recipient groups noted?		Typically, a write-up will include some analysis of baseline characteristics. A properly executed randomized control trial should result in baseline characteristics being balanced over the treatment and control groups, although chance imbalances can sometimes occur. Quasi-experimental studies typically have baseline imbalances, and so it is a feature of the quasi-experimental methods to try to address the taint that this may introduce.
If yes, for what variables was there imbalance?		
If yes, page, table, or figure number where this is reported.		
Did the study suffer from problems of unavailable data, non-response, or attrition?		
If yes, page, table, or figure number where this is discussed/reported.		
Pages, tables, or figures where baseline outcome data are presented		
Pages, tables, or figures where baseline sample characteristics are presented		

IDCG Risk of Bias Reporting (skip if study is not eligible for quantitative synthesis)		
Mechanism of assignment: was the allocation or identification mechanism able to control for selection bias?		
Group equivalence: was the method of analysis executed adequately to ensure comparability of groups throughout the study and prevent confounding?		
Hawthorne and John Henry effects: was the process of being observed causing motivation bias?		
Spill-overs: was the study adequately protected against performance bias?		
Selective outcome reporting: was the study free from outcome reporting bias?		
Selective analysis reporting: was the study free from analysis reporting bias?		
Other: was the study free from other sources of bias?		
Standard errors or confidence intervals: are appropriate standard errors or confidence intervals used?		
Quantitative Effect Estimates (skip if study is not eligible for quantitative synthesis)		
Deforestation		
Were effects reported for this outcome?		By "reported" we mean that they either appear clearly in some kind of table or graph or they are discussed in the text with some indication of the nature of the estimated effects. For example, to save space authors will sometimes omit tables or figures for effect estimates that are statistically insignificant but they will indicate that such estimates were constructed and that nothing significant was found. We want to be sure to include these in our list of reported outcomes so as not to bias our analysis by only including "significant" effects.
If yes...		
...what indicators were used?		List the indicators, separating by semi-colon (e.g., "forest cover percentage; annual deforestation rate", etc.).
...page, table, or figure number where the effects are reported:		Provide adequate reference so that someone can locate the effect estimate easily.
...where the effects found to be mostly positive, mostly negative, essentially zero, or mixed?		Of course if only one indicator is used, summarize with respect to that indicator. If multiple indicators were used, try to summarize on this basis.
...were the effects found to be mostly statistically significant, mostly insignificant, or mixed?		Statistically significant effects are those what have a two-sided p-value of 0.10 or less. This is often indicated by "stars" or other embellishments in tables that report results. By "mixed" we mean that the study reports multiple estimates of this effect, some of which are statistically significant but at least half or most are not.
Income, consumption, or poverty		
Were effects reported for this outcome?		
If yes...		
...what indicators were used?		
...page, table, or figure number where the effects are reported:		
...where the effects found to be mostly positive, mostly negative, essentially zero, or mixed?		
...were the effects found to be mostly statistically significant, mostly insignificant, or mixed?		

Excluded Studies

In the table that follows, we list studies that were subject to full text search but were then excluded on the basis of substantive or methodological grounds. Studies are listed by first three authors, date of publication, and then reasons for exclusion.

Author 1	Author 2	Author 3	Publication Year	PES or DFIM Intervention?	Deforestation or Poverty Outcome?	Developing Country?	Methodology Explained?	Quant. Criteria Satisfied?	Qual. Criteria Satisfied?
OC, Ajayi	Jack Kelsey B.	Leimona Bera	2012	Yes	No				
Pagiola, Stefano	Agustin Arcenas	Gunars Platais	2004	No					
Pagiola, Stefano			2007	No					
Parks, Peter	Ian W. Hardie		1995	No					
Pasha, Rachman	Leimona Bera		2011	No					
Pattanayak, Subhrendu	Sven Wunder	Paul J. Ferraro	2010	No					
Pedroni, Lucio	Michael Dutschke	Charlotte Streck	2013	No					
Phillips, VD			2011	No					
Pirard, Romain	Raphael Bille	Thomas Sembres	2010	No					
Pirard, Romain			2011	No					
Pilavin, Catherine	Bruno Guay	Lucio Pedroni	2011	No					
Purras, Ina			2010	Yes	Unclear	Yes	No		
Quintero, M.	S. Wunder	R.D. Estrada	2009	Yes	Yes	Yes	Yes	No	No
Rahloa, Sobataolo	Brian Mantlana	Harald Winkler	2011	No					
Ramirez, Octavio	Carlos E. Carpio	Rosalba Ortiz	2001	No					
Rowcroft, Pehina			2005	No					
Rutenbeck, Herman			1990	Unclear	Unclear	Yes	Unclear		
Sand, Isabel			2012	No					
Sandker, Maneke	Manuel Ruiz-Perz	Bruce M. Campbell	2012	No					
Sandker, Maneke	Samuel Kofi Nyame	Johannes Forster	2009	Unclear	No				
Seeborg Elverfeldt, Christina	Stefan Schwarze	Manfred Zeller	2008	Yes	Unclear	Yes	Yes		
Seeborg Elverfeldt, Christina	Stefan Schwarze	Manfred Zeller	2008	Yes	Yes	Yes	Yes	No	No
Slutsch, Margareta	Ben Vickers	Yola Georgiadou	2011	No					
Smith, Joyotee	Grahame Applegate		2002	No					
Spiteri, Anan	Sanjay K. Nepal		2006	No					
Subak, Susan			2000	No					
Suyanto, S.	Noviana Khususiyah	Bena Leimona	2007	No					
Swallow, Brent	Mikkel E. Kallesoe	Usman A. Iftikhar	2009	No					
Swallow, Brent	Bena Leimona	Thomas Yalich	2010	No					
Swallow, Brent	Ruth Meinzen-Dick	Meine van Noordwijk	2005	No					
Swallow, Brent	Ruth Meinzen-Dick		2009	No					
Tacconi, Luca			2009	No					
The, Bui	Hong Bich Ngoc		2006	Yes	Yes	Yes	Yes	No	No
Thuy, Pham	Bruce M. Campbell	Stephen Garnett	2009	No					
Thuy, Pham	Hoang Minh Ha	Bruce M. Campbell	2008	No					
Uchida, Emi	Scott Rozelle	Jinlao Xu	2009	Yes	No				
Viana, Virgilio	Maryanne Grieg-Gran	Rosana Della Misa	2009	No					
Villamor, Grace	Meine van Noordwijk		2011	No					
Villamor, G.B.	R.V. Cruz	R.D. Lasco	2006	Yes	No				
Wang, Xuejun	Jian Li		1999	No					
Wendland, Kelly	Lisa Naughton	Luis Suarez	2010	Yes	Yes	Yes	No		
Wendland, Kelly	Miroslav Horzák	Rosimery Portela	2009	No					
Wuenscher, Tobias	Stefanie Engel	Sven Wunder	2006	No					
Wuenscher, Tobias	Stefanie Engel	Sven Wunder	2008	No					
Wuenscher, Tobias	Stefanie Engel		2012	No					
Wunder, Sven			2006	No					
Wunder, Sven			n.d.	No					
Zahabu, E.	R.E. Malimbwi	Y.M. Ngaga	2005	No					
Zhang, Wei	Stefano Pagiola		2009	No					
Zilberman, David	Leslie Lipper	Nancy McCarthy	2008	Yes	Yes	No			
Yashoda	B.V. Chinnappa Reddy		2012	No					

Detailed risk of bias assessments

Program	Studies	Treatment Group	Method of attribution	Research design works with a source of exogenous assignment	Analysis controls for potential confounding due to land quality, socio-economic conditions, and accessibility	Measurement strategy avoids motivation bias	Study design accounts for spillover	Apparently free of selective outcome reporting	Apparently free of analysis fishing	Appropriate Statistical inference (standard errors)
China SLCP	Liu et al. 2010	2001-2004 SLCP households	multiple regression	No	No	Yes	No	Yes	Yes	Unclear
Costa Rica PSA	Arriagada et al. 2012	1997-2005 PSA conservation contract farms in Sarapiquí	matching and multiple regression	No	Yes	Yes	Yes	Yes	Yes	Yes
Costa Rica PSA	Arriagada et al. 2011	1998-2005 census tracts covered by PSA conservation contracts and on the common support with matched non-PSA tracts	matching	No	Yes	Yes	Unclear	Yes	Yes	Yes
Costa Rica PSA	Robalino and Pfaff (2013)	1997-2000 1kmX1km forested PSA conservation contracts pixels	matching and multiple regression	No	Yes	Yes	Yes	Yes	Yes	Yes
Costa Rica PSA	Robalino et al. 2008	2000-2005 1kmX1km forested PSA conservation contracts pixels	matching and multiple regression	No	Yes	Yes	Yes	Yes	Yes	Yes
Costa Rica PSA	Sanchez-Azofeifa et al.	1997-2000 grid cells covered by PSA	multiple regression	No	No	Yes	Unclear	Yes	Yes	Unclear

	2007	conservation contract								
Costa Rica RISE MP	Garbach et al. 2012	2002-2008 small/medium cattle farm properties in Esparza covered by PSA	multiple regression	No	No	No	No	Yes	Yes	Unclear
Mexico PSAH	Alix-Garcia et al. 2012	2004-2006 PSAH properties	matching and multiple regression	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Mexico PSAH	Scullion et al. 2011	2000-2003 PSAH or FIDECOAG UA properties	difference-in-differences	No	No	Yes	Unclear	Yes	Yes	No
Mexico MBCF	Honey-Roses et al. 2011	2000-2009 MBCF properties in Monarch Butterfly Biosphere Reserve area	matching and multiple regression	No	Unclear	Yes	Yes	Yes	Yes	Yes
Mozambique NCCL	Hegde and Bull 2011	2003-2006 NCCL participant households in project region	matching	No	Yes	No	No	Yes	Yes	Yes

Detailed Effect Estimates

PES and Forest Conversion

Program	Studies	Treatment Group	Time Period	Counterfactual Approximation	Forest Conservation Outcome	Forest Conservation Effect (standard error in parentheses)	ra-rc (standard error in parentheses)	ra
Costa Rica PSA	Arriagada et al. 2012	1997-2005 PSA conservation contract farms in Sarapiquí	1997-2005	(mean of estimates)	hectares of forest cover	8.73 (3.02)	0.012 (0.004)	0.010
				matched and difference-in-difference adjusted eligible farms in Sarapiquí without PSA	hectares of forest cover	12.09 (4.89)	0.017 (0.007)	0.015
				matched and regression adjusted farms in Sarapiquí without PSA (short cov set)	hectares of forest cover	9.70 (2.96)	0.014 (0.004)	0.012
				matched and regression adjusted farms in Sarapiquí without PSA (long cov set)	hectares of forest cover	8.48 (2.8)	0.012 (0.004)	0.010
				caliper matched and difference-in-difference adjusted eligible farms in Sarapiquí without PSA	hectares of forest cover	4.65 (1.41)	0.007 (0.002)	0.005
Costa Rica PSA	Arriagada et al. 2011	1998-2005 census tracts covered by PSA conservation contracts and on the common support with matched non-PSA tracts	1998-2005	nearest-neighbor matched census tracts without PSA (best covariate balance estimate)	hectares of forest deforested	-7.075 (3.91)	0.001 (0.001)	-0.011
					hectares of forest cover	29.443 (11.43)	0.005 (0.002)	0.010
Costa Rica PSA	Robalino and Pfaff (2013)	1997-2000 1kmX1km forested PSA conservation contracts pixels	1997-2000	(mean of estimates)	percent of forest area deforested	-0.335 (0.18)	0.001 (0.001)	-0.002

				bias-adjusted six-to-one propensity score matched forested pixels without PSA	percent of forest area deforested	-0.42	(0.15)	0.001	(0.001)	-	0.002
				bias-adjusted six-to-one covariate matched forested pixels without PSA	percent of forest area deforested	-0.44	(0.29)	0.001	(0.001)	-	0.002
				bias-adjusted six-to-one propensity score matched forested pixels without PSA, restricted to common support	percent of forest area deforested	-0.13	(0.09)	0.000	(0.000)	-	0.002
				bias-adjusted six-to-one covariate matched forested pixels without PSA, restricted to common support	percent of forest area deforested	-0.35	(0.29)	0.001	(0.001)	-	0.002
Costa Rica PSA	Robalino et al. 2008	2000-2005 1kmX1km forested PSA conservation contracts pixels	2000-2005	(mean of estimates)	percent of forest area deforested	-1.90	(0.7)	0.004	(0.002)	-	0.011
				bias-adjusted propensity score matched forested pixels without PSA	percent of forest area deforested	-1.87	(0.85)	0.004	(0.002)	-	0.011
				bias-adjusted covariate matched forested pixels without PSA	percent of forest area deforested	-1.92	(0.55)	0.004	(0.001)	-	0.011
Costa Rica PSA	Sanchez-Azofeifa et al. 2007	1997-2000 grid cells covered by PSA conservation contract	1997-2000	regression adjusted grid cells without PSA	probability that area was deforested	-0.002	(0.01)	0.001	(0.002)	-	0.003
Costa Rica RISEMP	Garbach et al. 2012	2002-2008 small/medium cattle farm properties in Esparza covered by PSA	2002-2008	convenience sample of local non-PES farmers	number of silvopastoral practices adopted	0.18	(0.37)				
		2002-2008 small/medium cattle farm properties in Esparza covered by	2002-2008	convenience sample of local non-PES farmers	number of silvopastoral practices adopted	0.79	(0.21)				

		PSA and technical assistance								
Mexico PSAH	Alix-Garcia et al. 2012	2004-2006 PSAH properties	2004-2006	(mean of estimates)	percent of forest area deforested	-1.29	(0.38)	0.007	(0.003)	-0.005
				Mahalanobis distance matched properties rejected in 2004 or online for PSAH for 2006 (best 90% matches)	percent of forest area deforested	-1.10	(0.35)	0.006	(0.003)	-0.005
				Mahalanobis distance matched and Tobit regression adjusted properties rejected in 2004 or online for PSAH for 2006 (best 90% matches)	percent of forest area deforested	-1.19	(0.34)	0.006	(0.003)	-0.005
				Inverse standard error distance matched properties rejected in 2004 or online for PSAH for 2006 (best 90% matches)	percent of forest area deforested	-1.57	(0.44)	0.008	(0.003)	-0.005
Mexico PSAH	Scullion et al. 2011	2000-2003 PSAH or FIDECOAGUA properties	2000-2003	local properties not under PSAH or FIDECOAGUA	hectares of pine oak forest cover	57	(na)	0.084	(na)	-0.004
					hectares of cloud forest cover	264	(na)	0.100	(na)	-0.031
Mexico MBCF	Honey-Roses et al. 2011	2000-2009 MBCF properties in Monarch Butterfly Biosphere Reserve area	2000-2009	spatial matched properties without MBCF	percent of forest area still covered by forest	2.60	(2.10)	0.003	(0.003)	-0.010
					percent of forest area undisturbed	11.60	(5.00)	0.016	(0.008)	-0.018

PES and welfare and estimated income effects

Program	Studies	Treatment Group	Time Period	Counterfactual Approximation	Welfare Outcome	Welfare Effect (standard error in parentheses)		Percentage Effect (standard error in parentheses)		
China SCLP	Liu et al. 2010	2001-2004 SLCP households	2001-2004	regression adjusted households without SLCP	Natural log of total household income	0.13	(0.03)	14%	(3.42)	
Mozambique NCCL	Hegde and Bull 2011	2003-2006 NCCL participant households in project region	2003-2006	(mean of estimates)	Consumption expenditure per capita in MTS	70039	(28778)			
				nearest neighbor matched local non-participant households stratification	Consumption expenditure per capita in MTS	89784	(34800)			
				matched local non-participant households radius	Consumption expenditure per capita in MTS	70762	(27427)			
				matched local non-participant households kernel	Consumption expenditure per capita in MTS	66409	(25740)			
				matched local non-participant households (mean of estimates)	Consumption expenditure per capita in MTS	53202	(27144)			
				(poor households)	Consumption expenditure per capita in MTS	13933	(23425)			
				(poor households)	nearest neighbor matched local non-participant households stratification	Consumption expenditure per capita in MTS	-3524	(5260)		
				(poor households)	matched local non-participant households radius	Consumption expenditure per capita in MTS	32896	(49099)		
				(poor households)	matched local non-participant households kernel	Consumption expenditure per capita in MTS	16566	(24725)		
				(poor households)	matched local non-participant households	Consumption expenditure per capita in MTS	9793	(14616)		

			(mean of estimates)	Household cash income per capita in MTS	278750	(108043)	4%	(1.55)
			nearest neighbor matched local non-participant households	Household cash income per capita in MTS	323000	(125194)	5%	(1.8)
			stratification matched local non-participant households	Household cash income per capita in MTS	249000	(96511)	4%	(1.38)
			radius matched local non-participant households	Household cash income per capita in MTS	277000	(107364)	4%	(1.54)
			kernel matched local non-participant households	Household cash income per capita in MTS	266000	(103101)	4%	(1.48)
	(poor households)		(mean of estimates)	Household cash income per capita in MTS	103256	(154113)	18%	(38.13)
	(poor households)		nearest neighbor matched local non-participant households	Household cash income per capita in MTS	53024	(79140)	8%	(14.98)
	(poor households)		stratification matched local non-participant households	Household cash income per capita in MTS	135000	(201493)	25%	(60.28)
	(poor households)		radius matched local non-participant households	Household cash income per capita in MTS	116000	(173134)	21%	(46.13)
	(poor households)		kernel matched local non-participant households	Household cash income per capita in MTS	109000	(162687)	19%	(41.61)

International Initiative for Impact Evaluation
London International Development Centre
36 Gordon Square
London WC1H 0PD
United Kingdom

3ieuk@3ieimpact.org
Tel: +44 207 958 8351/8350



www.3ieimpact.org