Policy Perspective, Conservation Letters

Mainstreaming impact evaluation in nature conservation

Kathy Baylis[1], Jordi Honey-Rosés[2], Jan Börner[3], Esteve Corbera[4], Driss Ezzine-de-Blas [5], Paul Ferraro [6], Renaud Lapeyre [7], U. Martin Persson[8], Alex Pfaff [9], & Sven Wunder[10]

1 Agriculture and Consumer Economics, University of Illinois Urbana-Champaign, Champaign, IL 61820, U.S.A baylis@illinois.edu
2 School of Community and Regional Planning, University of British Columbia, Vancouver, B.C. V6T 1Z2, Canada jhoney@mail.ubc.ca
3 Center for Development Research, University of Bonn & Center for International Forestry Research (CIFOR) Bonn, Germany jborner@uni-bonn.de
4 Institute of Environmental Science and Technology (ICTA), Autonomous University of Barcelona, Bellaterra, Catalonia, Spain esteve.corbera@uab.cat
5 Centre de coopération internationale en recherche agronomique pour le développement (CIRAD), Montpellier, France ezzine@cirad.fr
6 Department of Economics, Andrew Young School of Policy Studies, Georgia State University, Atlanta Georgia 30302, U.S.A pferraro@gsu.edu
7 Institut du développement durable et des relations internationales (IDDRI), Paris, France renaud.lapeyre@iddri.org
8 Chalmers University of Technology, Sweden martin.persson@chalmers.se
9 Sanford School of Public Policy, Duke University, U.S.A. alex.pfaff@duke.edu
10 Center for International Forestry Research (CIFOR), Embrapa Amazônia Oriental – Convênio CIFOR, Trav. Dr. Enéas Pinheiro s/n, CEP 66.095-100 Belém -PA, Brazil. s.wunder@cgiar.org

Short Running Title: Mainstreaming Impact Evaluation

Keywords: Biodiversity, Conservation Policy, Impact Evaluation, Payment for Environmental Services, Protected Areas

Word count abstract:	118
Word count manuscript:	4363
N references:	52
N figures & tables:	0

Corresponding Author: Jan Börner. Center for Development Research, University of Bonn. Bonn, Germany jborner@uni-bonn.de +49-228-73-1873

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1111/conl.12180.

Abstract

An important part of conservation practice is the empirical evaluation of program and policy impacts. Understanding why conservation programs succeed or fail is essential for designing cost-effective initiatives and for improving the livelihoods of natural resource users. The evidence we seek can be generated with modern impact evaluation designs. Such designs measure causal effects of specific interventions by comparing outcomes with the interventions to outcomes in credible counterfactual scenarios. Good designs also identify the conditions under which the causal effect arises. Despite a critical need for empirical evidence, conservation science has been slow to adopt these impact evaluation designs. We identify reasons for the slow rate of adoption, and provide suggestions for mainstreaming impact evaluation in nature conservation.

Introduction

Conservation science is only slowly beginning to build a body of evidence on the impact of conservation policies (Ferraro & Pattanayak 2006; Fisher et al. 2014). Many compelling reasons motivate impact evaluations of conservation policy instruments. Organizations want to know where to invest scarce resources, while governments and donors seek tangible outcomes. Evidence of why conservation initiatives succeed or fail is also essential for designing cost-effective programs and improving the livelihoods of natural resource users (Sutherland et al. 2004; Cook et al. 2010). In this paper, we propose steps toward mainstreaming and improving conservation policy impact evaluation.

Impact evaluation has developed into a research discipline with multiple fields of application including health, education and development (White 2009). Our notion of impact evaluation goes beyond monitoring program inputs, outputs or indicators over time. It measures the causal effect of a specific policy, program or intervention vis-à-vis a credible counterfactual scenario and seeks to understand the conditions under which this effect arises (Ferraro & Hanauer 2014). In a comprehensive impact evaluation, evaluators will rule out alternative or rival explanations of program outcomes (Ferraro 2009). One might also examine past outcomes to forecast the potential impact of future interventions (Pfaff et al. 2009). To obtain these insights, impact evaluations must be more than abstract quantitative evaluations but rather build on qualitative theories of change which help identify the conditions in which the desired impacts arise (Morgan & Winship 2007).

We are not the first to make the points above. Several recent papers call for improving the quality of impact evaluation in nature conservation (Ferraro & Pattanayak 2006; Miteva et al. 2012; Pullin 2012; Fisher et al. 2014). Despite these calls, conservation science still lags behind health, education, and development policy in adopting best practices in impact evaluation (Banerjee & Duflo 2009). Few studies meet even the basic standards of an impact evaluation such as considering before and after conditions, including control groups, accounting for confounding factors, or systematically ruling out rival hypotheses (Bowler et al. 2012; Samii et al. 2014).

In contrast to earlier essays on this subject, we explore the reasons why nature conservation policy has been slow to adopt more rigorous impact evaluation designs. The reasons are not trivial and the solutions are not simple. We characterize the current barriers and propose elements of a strategy that may build a systematic body of evidence on the effectiveness of conservation initiatives. Our arguments are based on discussions from the workshop "*Evaluating Forest Conservation Initiatives: New Tools and Policy Needs*" organized in Barcelona, Spain in December 2013.

Challenges for Impact Evaluation in Conservation Science

Conservation programs have features that, while not unique to conservation, translate into specific challenges for impact evaluation.

(1) **Multiple outcomes and scales**. Conservation interventions often strive to achieve multiple objectives at multiple scales. For instance, ensuring viable species populations while protecting habitat; or maintaining ecosystem integrity while increasing the provision of ecosystem services for human populations. 'Co-benefits' may be relevant in other contexts, but in conservation, co-benefits are often central to program success. The backlash against the Reduced Emissions from Deforestation and Forest Degradation (REDD) program, for initially focusing only on carbon capture as a singular metric illustrates the distaste for single policy objectives in a multiple-output setting (Corbera et al. 2010). Furthermore, ecosystems are complex systems with non-linear dynamics at various spatial and temporal scales (Fisher et al. 2009; Koch et al. 2009). Such complexity raises practical hurdles. For example, the conservation of migratory species requires management in both breeding and wintering grounds, often distributed across multiple ecosystems and political administrations (Brower 1995; Naidoo et al. 2014). In the presence of multiple objectives at multiple scales it also becomes more difficult to articulate clear theories of change and empirical strategies for impact evaluation. Different choices about scale will also inhibit comparable replications.

(2) **Spatial spillovers**. While many fields can ignore the spatial component of an impact evaluation, conservation simply cannot. Space is an essential part of ecological processes: water flows, pollution emissions, species migration, deforestation, and dispersal. Therefore to assess the impact of conservation policies one must account for the appropriate spatial scale. Yet even when the appropriate spatial scale is well known, measuring the net impacts of an intervention is complicated by spatial spillovers. These spillovers can be a result of ecological process, but can also result from behavioral responses, such as when restricting access to resources in one area induces a rise in extractive activity elsewhere, in what is referred to as "leakage" (Ostwald & Henders 2014). Spillovers not only affect net impacts but can also bias impact estimation when they influence non-target areas that were intended to serve as control observations.

(3) **Confounding factors**. Many biophysical, behavioral and institutional factors affect both where conservation initiatives take place and the outcomes we measure. Imagine that the survival of a particular species depends on forest habitat under threat by logging pressures. Policy makers respond by creating a new protected area, but the location and boundaries of protection are developed in consultation with local municipalities who prioritize remote areas far from human settlements. An impact evaluation that were to compare conservation outcomes inside this park with conservation outcomes outside the park might erroneously find that the park was highly successful if areas with low deforestation risk were protected, while areas with easier access, closer to human settlement, and high deforestation risk were left unprotected (Joppa & Pfaff 2010). Assume further that timber values increased after the park was created, resulting in a generalized spike in logging. A before-after comparison might lead to the erroneous conclusion that the park was unsuccessful. In both cases, these approaches fail to address the confounding factors affecting protected area placement and outcome. These confounding factors must be accounted for in non-experimental evaluations. To do so, evaluators need to draw on expertise

from various disciplines and on-the-ground knowledge. In cases in which confounders are not easily observable, evaluators often use instrumental variables – variables that only affect the outcome through their effects on the probability of participating in a program (e.g., weather conditions or other shocks like natural disasters). Finding such variables in the conservation context is difficult because they often affect conservation outcomes directly.

(4) **Randomization's limits**. Conservation science has been slow to adopt randomized controlled trials (RCT). Notable exceptions include Ferraro et al. (2011), Jack (2013), Samii et al. (2014) and experiments in habitat and invasive species management studies (Sutherland et al. 2004). Practical and ethical considerations often limit the successful use of RCTs. Randomization is not viable with small sample sizes or low replication, and research designs must be adjusted accordingly. However, it is difficult to obtain large sample sizes or replicate if a single intervention covers a large geographic area. For example, to randomize a program to preserve water quality with acceptable statistical power, one might have to treat hundreds of watersheds. RCTs also rely on the "stable unit value treatment assumption", which implies outcomes in one observation are not affected by the treatment status of another. This assumption may not hold in the presence of spatial spillovers: where the outcome of one parcel affects neighboring parcels. Last, it may also be politically untenable or unethical to randomly distribute restrictive conservation regulations. Conservation policy evaluation will thus have to rely on paired research designs and innovative quasi-experimental approaches, such as regression discontinuity design and synthetic control analysis (e.g., Abadie et al. 2014).

(5) **Small initiatives**. Large-scale and generously funded pilot initiatives are rare in the conservation sector, constraining even those firmly committed to measuring impact. Innovative program designs are often developed by small organizations that integrate multiple funding streams and gradually develop their intervention design through years of experience. For these organizations, it may not be feasible to embark on impact evaluation by themselves. Either outside support or some critical mass of similar interventions is probably needed to carry out full-scale evaluation designs.

Implications for conservation policy and science

Several implications arise for the design of impact evaluations and the effective integration of evaluation results in a "conservation policy research cycle", where the knowledge base is continuously updated as new evidence emerges. First, to refine theories of change, researchers need to **cross epistemological divides and integrate qualitative and quantitative approaches** (Margoluis et al. 2009; Agrawal & Chhatre 2011). Qualitative understanding helps contextualize quantitative treatment effect estimates and quantitative methods can inform qualitative research design and theory development. As an example consider the use of quantitative data to inform the selection of locations for in-depth qualitative analysis, either by targeting outliers or more representative sampling groups (Roe et al. 2013). Multi-disciplinary perspectives should not only inform theories of change and related intervention designs, they can also help to develop more appropriate evaluation strategies (White 2009).

The second implication relates to **choosing the appropriate scale of analysis**. In conservation practice, the unit of analysis, spatial scale and outcome variable is not always readily apparent. As a starting point, the analytical scale should be motivated by the theory of change. And yet the decision-making unit and the resource governance regime in conservation programs can be nebulous: ecosystems are co-managed by private owners, collectives, communities or state agents. The issue is further complicated because the natural, social and medical sciences differ in their view on what constitutes the 'right' scale, unit of analysis, and appropriate sample size. Since social and ecological processes operate at multiple scales, any single choice of scale will inevitably fail to capture certain dynamics. Researchers often compare conservation outcomes across a landscape by dividing their study areas into a uniform grid. However uniform grids inevitably combine multiple ecosystem types, governance regimes or property owners. Thus the choice of analytical unit, as well as geographical (e.g. valley, watershed, landscape, ecosystem) and administrative (e.g. community, municipality, county, region) scale is challenged by methodological constraints. The solution is not found in selecting fine-grained analytical units because high resolutions will generate spatial correlations that bias results. Conversely, coarse resolutions may fail to capture local processes and inhibit the identification of appropriate controls. Selecting the appropriate scale can reduce the unobserved confounding factors, while using an inappropriate scale can exacerbate the effects of unobserved confounding factors. Where the appropriate scale is unknown, impact evaluations may use hierarchical models or replicate the analysis at multiple scales to evaluate how sensitive results are to the choice of scale (eg Avelino et al. 2015; Börner et al. 2015; Costedoat et al. 2015).

The third implication relates to **incorporating spillover effects into the research design**, including leakage, spatial autocorrelation and peer effects. Several recent papers explore the effect of conservation policies on conservation outcomes in neighboring areas. Some report that an increase in protection in one area displaces deforestation activities to other areas (Oliveira et al. 2007; Meyfroidt & Lambin 2009). Others find a “halo effect” whereby areas adjacent to protected areas are better protected than one might expect (Honey-Rosés et al. 2011; Gaveau et al. 2012; Robalino & Pfaff 2012). Ideally, theories and related evaluation methods would address both potential sources of spillovers, behavioral and mechanistic, since the existence of either should be part of the estimated treatment effect. Ignoring spillover effects will bias estimates of program impact.

Fourth, while **randomization** might not be possible for programs that require large, contiguous areas, some conservation instruments, such as Payments for Ecosystem Services (PES) or community-based programs are amenable to randomization, particularly if the desired environmental outcomes are local. For example, incentive based contracts are being randomly allocated in the mountains of Bolivia (Asquith et al. 2008; Jones 2012) and Uganda (Hatanga 2014). Where feasible and ethical, randomizing treatment can help researchers address potential confounding factors by ensuring they are not associated with treatment. Randomization may also be used when a program is thought to work, and program managers would like to test variations of the program or specific aspects of its mechanism to identify why and how the program produces the desired results. It also might be easier to randomize over enforcement than over the placement of protected areas, or it might be possible to randomize over the type of PES contract needed to induce changes in household behavior. Further, if a country is interested in introducing a nation-wide conservation effort, randomizing over location might be feasible. However, any randomized intervention would require an important investment in communicating its purpose and the targeting rationale because, as noted earlier, such approaches may entail political and ethical challenges. Further, randomization needs to be part of a broader evaluation strategy that incorporates qualitative work to explore the causal chain.

Finally, impact evaluation in conservation should be **sensitive to heterogeneous outcomes** (Alix-Garcia et al. 2012; Pfaff & Robalino 2012; Ferraro & Miranda 2013). Conservation policies and programs affect a variety of social actors under varying bio-physical conditions. Moving beyond the average effects of an intervention, conservation planners need to know where and for whom it worked (Deaton 2009). Estimating heterogeneous treatment effects and uncovering causal mechanisms behind average treatment effects is difficult and can increase the complexity of the research design. It requires even more elaborate theories and more untestable (and often less credible) assumptions than are required to estimate unconditional effects. Thus, even in randomized controlled trials, estimates of heterogeneous treatment effects are considered much less credible than unconditional effects unless incorporated directly in the experimental design (Ferraro & Hanauer 2014).

Moving Forward

Building a body of evidence on conservation policy effectiveness will require greater collaboration between researchers and conservation managers akin to the long-standing partnership between medical scholars and clinicians. The evidence-base we advocate for is a global public good. Therefore it is not surprising that it has been difficult to muster the resources necessary for building a solid evidence-base. Unless practitioners are strongly encouraged by donors, we will end up with an under-provision of evidence; i.e., the status quo.

Systematic reviews and systematic maps (or evidence gap maps) are a useful tool to synthesize scientific results and identify shortfalls for policy makers (Dicks et al. 2014). The Initiative for Impact Evaluation (www.3ieimpact.org) and the Collaboration for Environmental Evidence (www.environmentalevidence.org) have recently published systematic reviews of the effect of protected areas, payment for ecosystem services and aspects of forest management on various human welfare, habitat and species preservation outcomes. In all of these reviews, authors point to the limited or fragmented evidence of the effect of these various policy instruments (Bowler et al. 2010; Geldmann et al. 2013; Pullin et al. 2013; Samii et al. 2014). Protected areas have arguably received the most attention with the reviews analyzing 86 articles on habitat and species outcomes and 306 articles on perceptions of PAs and 79 on welfare impacts. Nevertheless, large gaps remain. Even for protected areas, our understanding of spillover and heterogeneous treatment effects on environmental and socioeconomic indicators is still limited (Pullin et al. 2013; Geldman et al. 2013). Empirical evidence is even more sparse on the effectiveness and implementation costs of other large and small scale conservation policy instruments, such as forest law enforcement, PES schemes, eco-certification, and Integrated Conservation and Development (ICDP) approaches that dominate public and private REDD+ initiatives (Blom et al. 2010; Lambin et al. 2014). Furthermore, we need to know how these policies compare in their effect on human and environmental outcomes, and how these instruments work in policy mixes (Barton et al. 2013).

Mainstreaming impact evaluation in conservation will require partnerships between scientists and program implementers during the design phase to: (1) clarify program objectives, possibly with a modification in design; (2) identify a theory of change, counterfactual groups, and testable hypotheses; and (3) define performance indicators and data collection protocols. Wherever feasible, a randomized program design may reduce rather than increase the costs of impact evaluation, particularly in sub-national or single project pilot interventions. Such partnerships should be maintained over time to facilitate continuous feedbacks between evaluation, design protocols and criteria, and implementation practice, which all should be flexible, adaptive and responsive to assessment outcomes (Sims et al. 2014). Donors could support this process by conditioning funding, including performance bonuses on well-designed impact evaluation, and collaborating with researchers on defining priorities for focused and carefully designed systematic reviews (CEE 2013). Most importantly, impact evaluation needs long-term support to build a strategic global evidence base for conservation policies (Keene & Pullin 2011). Such growing collaboration and mutual understanding should slowly preclude awarding funding on the basis of exaggerated *ex-ante* claims of conservation potential, since this is counterproductive to building a solid body of evidence. Conservation at the scale envisaged by international policy initiatives, such as REDD+, clearly stands to benefit from a solid body of evidence on what works, what does not, where and why.

Acknowledgements

The perspectives put forward in this paper are based on discussions among scientists and practitioners who attended the workshop “Evaluating Forest Conservation Initiatives: New Tools and Policy Needs” on 10-12 December 2013, in Barcelona, Spain. We are grateful for support from the European Association of Environmental and Resource Economists, the Robert Bosch Foundation, the Autonomous University of Barcelona, the Center for International Forestry Research, and the ERA-Net BiodivERsA funded INVALUABLE project with the French national funder Agence Nationale de la Recherche. The INVALUABLE project generously covered publication fees. We also thank the workshop participants: A. Angelsen, R. Arriagada, N. Asquith, M. Bottrill, J. Bush, E. Cisneros, S. Costedoat, P. Davies, S. Desbureaux, L. Glew, E. Górriz, B. Hedden-Dunkhorst, N. Hogarth, C. Loucks, A. Leonardi, B. Macura, A. Martin, M. Masozera, D. Miteva, S. Pagiola, V. Rodríguez Osuna, J. Schielein, P. Shah, E. Sills, J.M. Torres, E. Varela, G. Le Velly, C. Xie, and S. L. Zhou. Finally we thank three anonymous reviewers who improved the manuscript.

References

Abadie, A., Diamond, A. & Hainmueller, J. (2014). Comparative Politics and the Synthetic Control Method. *American Journal of Political Science*, doi:10.1111/ajps.12116.

Agrawal, A. & Chhatre, A. (2011). Strengthening causal inference through qualitative analysis of regression residuals: Explaining forest governance in the Indian Himalaya. *Environment and Planning A*, 43, 328–346.

Alix-Garcia, J.M., 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, 613–638.

Asquith, N.M., Teresa, M. & Wunder, S. (2008). Selling two environmental services: In-kind payments for bird habitat and watershed protection in Los Negros, Bolivia. *Ecological Economics*, 65, 675–684.

Avelino, A.F.T., Baylis, K. & Honey-Rosés, J. (2015). *Goldilocks and the grid: Selecting scale when evaluating conservation polities*. *University of Illinois Working Paper*.

Banerjee, A.V. & Duflo, E. (2009). The Experimental Approaches to Development Economics. *Annual Review of Economics*, 1, 151–178.

Barton, D.N., Blumentrath, S. & Rusch, G. (2013). Policyscape—A Spatially Explicit Evaluation of Voluntary Conservation in a Policy Mix for Biodiversity Conservation in Norway. *Society & Natural Resources*, 26, 1185–1201.

Blom, B., Sunderland, T. & Murdiyarso, D. (2010). Getting REDD to work locally: lessons learned from integrated conservation and development projects. *Environmental Science & Policy*, 13, 164–172.

Börner, J., Kis-Katos, K., Hargrave, J. & König, K. (2015). Post-crackdown effectiveness of field-based enforcement in the Brazilian Amazon. *PloS One*, (in press).

Bowler, D., Buyung-Ali, L.M., Healey, J.R., Jones, J.P.G., Knight, T.M. & Pullin, A.S. (2010). *The Evidence Base for Community Forest Management as a Mechanism for Supplying Global Environmental Benefits and Improving Local Welfare*.

Bowler, D.E., Buyung-Ali, L.M., Healey, J.R., Jones, J.P., Knight, T.M. & Pullin, A.S. (2012). Does community forest management provide global environmental benefits and improve local welfare? *Frontiers in Ecology and the Environment*, 10, 29–36.

Brower, L.P. (1995). Understanding and misunderstanding the migration of the monarch butterfly (Nymphalidae) in North America: 1857-1995. *Journal of the Lepidopterists Society*, 49, 304–385.

CEE. (2013). *Guidelines for Systematic Review and Evidence Synthesis in Environmental Management*. *Collaboration for Environmental Evidence*. www.environmentalevidence.org/Documents/Guidelines/Guidelines4.2.pdf.

Cook, C.N., Hockings, M. & Carter, R. (2010). Conservation in the dark? The information used to support management decisions. *Frontiers in Ecology and the Environment*, 8, 181–186.

Corbera, E., Estrada, M. & Brown, K. (2010). Reducing greenhouse gas emissions from deforestation and forest degradation in developing countries: revisiting the assumptions. *Climatic Change*, 100, 355–388.

Costedoat, S., Corbera, E., Ezzine-de-Blas, D., Baylis, K., Honey-Rosés, J. & Castillo-Santiago, M.A. (2015). How effective are biodiversity payments in Mexico? *PloS ONE*. 10(3): e0119881 doi 10.1371/journal.pone.0119881

Deaton, A.S. (2009). Instruments of Development: Randomization in the tropics and the search for the elusive keys to economic development. *NBER Working Paper 14690*.

Dicks, L. V., Walsh, J.C. & Sutherland, W.J. (2014). Organising evidence for environmental management decisions: a "4S" hierarchy. *Trends in Ecology & Evolution*, 29, 607–613.

Ferraro, P.J. (2009). Counterfactual Thinking and Impact Evaluation in Environmental Policy. *Environmental program and policy evaluation: Addressing methodological challenges. New Directions for Evaluation*, 75–84.

Ferraro, P.J. & Hanauer, M.M. (2014). Advances in Measuring the Environmental and Social Impacts of Environmental Programs. *Annual Review of Environment and Resources*, 39, 495-517.

Ferraro, P.J. & Miranda, J.J. (2013). Heterogeneous treatment effects and mechanisms in information-based environmental policies: Evidence from a large-scale field experiment. *Resource and Energy Economics*, 35, 356–379.

Ferraro, P.J., Miranda, J.J. & Price, M.K. (2011). The Persistence of Treatment Effects with Norm-Based Policy Instruments: Evidence from a Randomized Environmental Policy Experiment. *American Economic Review*, 101, 318–322.

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

Fisher, B., Balmford, A., Ferraro, P.J., Glew, L., Mascia, M., Naidoo, R. & Ricketts, T.H. (2014). Moving Rio Forward and Avoiding 10 More Years with Little Evidence for Effective Conservation Policy. *Conservation Biology*, 28, 880–882.

Fisher, B., Turner, R.K. & Morling, P. (2009). Defining and classifying ecosystem services for decision making. *Ecological Economics*, 68, 643–653.

Gaveau, D.L.A., Curran, L.M., Paoli, G.D., Carlson, K.M., Wells, P. & Ratnasari, D. (2012). Examining protected area effectiveness in Sumatra: importance of regulations governing unprotected lands. *Conservation Letters*, 5, 142–148.

Geldmann, J., Barnes, M., Coad, L., Craigie, I., Hockings, M. & Burgess, N. (2013). Effectiveness of terrestrial protected areas in reducing biodiversity and habitat loss. *Collaboration for Environmental Evidence*.

Hatanga, P. (2014). *Random trials in Uganda to show if payments for ecosystem services really work*. *International Institute for Environment and Development*. http://www.iied.org/random-trials-uganda-show-if-payments-for-ecosystems-services-really-work.

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

Jack, B.K. (2013). Private Information and the Allocation of Land Use Subsidies in Malawi. *American Economic Journal: Applied Economics*, 5, 113–135.

Jones, J. (2012). Can Paying for Global Ecosystem Services reduce poverty? *Funded by Ecosystem Services for Poverty Alleviation Project*, http://www.espa.ac.uk/projects/espa–2012–grants.

Joppa, L. & Pfaff, A. (2010). Reassessing the forest impacts of protection: The challenge of nonrandom location and a corrective method. *Annals of the New York Academy of Sciences*, 1185, 135–149.

Keene, M. & Pullin, A.S. (2011). Realizing an effectiveness revolution in environmental management. *Journal of Environmental Management*, 92, 2130–5.

Koch, E.W., Barbier, E.B., Silliman, B.R., Reed, D.J., Perillo, G.M., Hacker, S.D., Granek, E.F., Primavera, J.H., Muthiga, N., Polasky, S., Halpern, B.S., Kennedy, C.J., Kappel, C. V, Wolanski, E. & Services, E. (2009). Non-linearity in ecosystem services: temporal and spatial variability in coastal protection. *Frontiers in Ecology and the Environment*, 7, 29–37.

Lambin, E.F., Meyfroidt, P., Rueda, X., Blackman, A., Börner, J., Cerutti, P.O., Dietsch, T., Jungmann, L., Lamarque, P., Lister, J., Walker, N.F. & Wunder, S. (2014). Effectiveness and synergies of policy instruments for land use governance in tropical regions. *Global Environmental Change*, 28, 129–140.

Margoluis, R., Stem, C. & Salafsky, N. (2009). Design Alternatives for Evaluating the Impact of Conservation Projects. *Environmental program and policy evaluation: Addressing methodological challenges. New Directions for Evaluation*, 85–96.

Meyfroidt, P. & Lambin, E.F. (2009). Forest transition in Vietnam and displacement of deforestation abroad. *Proceedings of the National Academy of Sciences of the United States of America*, 106, 16139–16144.

Miteva, D.A., Pattanayak, S.K. & Ferraro, P.J. (2012). Evaluation of biodiversity policy instruments: what works and what doesn't? *Oxford Review of Economic Policy*, 28, 69–92.

Morgan, S.L. & Winship, C. (2007). *Counterfactuals and causal inference: methods and principles for social research*. Cambridge University Press, Cambridge, UK.

Naidoo, R., Chase, M.J., Beytell, P., Du Preez, P., Landen, K., Stuart-Hill, G. & Taylor, R. (2014). A newly discovered wildlife migration in Namibia and Botswana is the longest in Africa. *Oryx*, 1–9.

Oliveira, P.J.C., Asner, G.P., Knapp, D.E., Almeyda, A., Galván-Gildemeister, R., Keene, S., Raybin, R.F. & Smith, R.C. (2007). Land-use allocation protects the Peruvian Amazon. *Science*, 317, 1233–1236.

Ostwald, M. & Henders, S. (2014). Making two parallel land-use sector debates meet: Carbon leakage and indirect land-use change. *Land Use Policy*, 36, 533–542.

Pfaff, A. & Robalino, J. (2012). Protecting forests, biodiversity, and the climate: predicting policy impact to improve policy choice. *Oxford Review of Economic Policy*, 28, 164–179.

Pfaff, A., Robalino, J., Sanchez-Azofeifa, G.A., Andam, K.S., Ferraro, P.J. (2009). Park Location Affects Forest Protection: Land Characteristics Cause Differences in Park Impacts across Costa Rica. *The B . E . Journal of Economic Analysis & Policy*, 9, Article 5.

Pullin, A.S. (2012). Realising the potential of environmental data: a call for systematic review and evidence synthesis in environmental management. *Environmental Evidence*, 1, 2.

Pullin, A.S., Bangpan, M., Dalrymple, S., Dickson, K., Haddaway, N.R., Healey, J.R., Hauari, H., Hockley, N., Jones, J.P.G., Knight, T., Vigurs, C. & Oliver, S. (2013). Human well-being impacts of terrestrial protected areas. *Environmental Evidence*, 2, 19.

Robalino, J.A. & Pfaff, A. (2012). Contagious development: Neighbor interactions in deforestation. *Journal of Development Economics*, 97, 427–436.

Roe, D., Grieg-Gran, M. & Mohammed, E.Y. (2013). Assessing the social impacts of conservation policies: rigour versus practicality. *IIED Briefing Papers. International Institute for Environment and Development*.

Samii, C., Lisiescki, M., Kulkarni, P., Paler, L. & Chavis, L. (2014). Effects of Decentralized Forest Management (DFM) on Deforestation and Poverty in Low and Middle Income Countries: A Systematic Review CEE13-015a. *Collaboration for Environmental Evidence*.

Sims, K.R.E., Alix-Garcia, J.M., Shapiro-Garza, E., Fine, L.R., Radeloff, V.C., Aronson, G., Castillo, S., Ramirez-Reyes, C. & Yañez-Pagans, P. (2014). Improving environmental and social targeting through adaptive management in Mexico's payments for hydrological services program. *Conservation Biology*, 28, 1151–9.

Sutherland, W.J., Pullin, A.S., Dolman, P.M. & Knight, T.M. (2004). The need for evidence-based conservation. *Trends in ecology & evolution*, 19, 305–8.

White, H. (2009). Theory-based impact evaluation: principles and practice. *Journal of Development Effectiveness*, 1, 271–284.