Appendix A: Supplementary methods

Literature review search protocol

Literature searches covered three online databases: Web of Science (WoS), CAB Abstracts (CAB) and Google Scholar (GS). Data from WoS and CAB were exported on April 22nd 2018, and explored first, and GS searches began on July 2nd 2019, and continued until November 10th 2019. Search terms were developed and applied for each of the 12 SDGs deemed to be 'non-environmental' (i.e. SDGs 1 to 11, and 16), and excluding targets in the remaining SDGs deemed environmental. We considered a target to be environmental if its wording suggests that achievement of the target rests wholly on conserving, protecting or improving some aspect of the natural environment. Searches were based on keywords taken from the official SDG targets and indicators[1]. Keywords were defined as any word or short phrase that applies specifically to at least one target within one or more goal. Ambiguous and grammatically irrelevant (e.g. articles, pronouns) words were omitted. To avoid excessive irrelevant returned items, words with multiple meanings or applications (e.g. 'health') were used only with associated words occurring in the SDG targets and indicators (e.g. 'health personnel', 'mental health', 'health-care service' etc.). All search terms were reviewed by co-authors, and passed through two thesauruses (the CAB Thesaurus, associated with the CAB Abstracts, and http://www.thesaurus.com) and relevant synonyms added as appropriate.

The search functions for GS differ from those used by the other two databases and so were conducted as separate component. For all searches, target-related search terms were coupled (i.e. using a Boolean 'and' clause) with standardised search terms used to target papers that make specific reference to forests or similar habitats (rainforest* OR forest* OR woodland* OR jungle* OR mangrove*), and for the WoS, CAB searches we included an additional search clause to target papers with a focus on external interventions (intervention* OR project* OR strateg* OR program* OR scheme* OR enterprise*), which we consider to the be the most insightful form of evidence relating to our topic. Searches in WoS used the topic field (TS), and those in CAB used the abstract, title and descriptor fields (ab, ti and de). We acknowledge that the use of an additional clause in two of our three searches will have limited some of our findings, but we considered this to be a reasonable way to limit search outputs to a more manageable number whilst giving priority to the evidence that we consider most robust.

Datasets from the WoS and CAB searches were combined and duplicates were removed using (i) unique identification numbers provided by the databases; (ii) DOI numbers; and (iii) a combination of title and year, the latter undergoing manual checks to avoid erroneous deletions. This process yielded a final list of 55,167 publications for review, although a few duplicate papers remained and were removed manually at a later stage.

GS searches used the advanced search option, excluding patents and citations, and placing the target-related term in the "with the exact phrase" section and the forest terms in the "with at least one of the words" section. For each target-related term, searches were run twice, first using the "in the title of the article" option and then using the "anywhere in the article option". For each search, the first 100 items listed were considered for inclusion. A total of 978 searches were conducted, although duplicate items were not monitored for this component, so the total number of unique items considered is not clear.

Inclusion/exclusion criteria

Applying the inclusion/exclusion criteria given in the following paragraphs, consideration of individual items was based on an established review protocol[2] of filtering by title, then by abstract, followed by extraction of information. This was conducted by three reviewers (JC, NT and JW-H), using checks for consistency (kappa analysis) between reviewers on randomly-selected subsets of 100 items. Kappa scores of 0.7 were used as the accepted threshold, and, where consistency checks produced values below this, reviewers discussed their choices and repeated the process (using a new subset) until a suitable score was achieved.

For WoS- and CAB-derived literature, basic inclusion criteria were that studies documented one or more external interventions aiming to achieve progress towards one or more non-environmental SDG targets, and used a forest-related measure as an outcome variable. Explicit mention of the SDGs or its targets was not required. For GS searches, inclusion was not limited to studies looking at specific interventions, and included any item that made reference to a target-related topic with some indication of expected/observed forest impacts.

As the WoS and CAB searches focused on intervention-related studies with observed impacts, in addition to the criteria listed below, which applied to all searches, we also excluded from these searches studies based solely on predictive modelling or speculative (i.e. non-empirical) outcomes, and cases where interventions sought to achieve multiple goals or targets at the same time, making identification of target-specific impacts not possible. This included large-scale resettlement/transmigration programmes, which, in addition to presenting problems of target-level impact identification, often fail to meet their development objectives[3], may be influenced by private sector interests (especially large agri-business)[4], and/or can result in human rights violations[5,6]. Nevertheless, we acknowledge the roles that such schemes play in forest dynamics and human development, and, in many cases, the recommendations made in our main article are still applicable.

Throughout all searches, the following criteria were applied:

- Publications focusing on payments for ecosystem services (PES), reducing emissions from deforestation and degradation (REDD) or 'alternative livelihoods' schemes, which were all deemed environmental, were excluded, unless explicit mention was given to a specific development component that did not depend upon the achievement of a forest- or conservation-related outcome for the scheme's success.
- Publications documenting community-based natural resource management and related schemes (e.g. participatory forest management, joint forest management etc.) were included, provided reference was made to an explicit development objective deemed independent and not reliant upon the achievement of a forest- or conservation-related outcome for its success.
- Publications in languages other than English were excluded, due to a lack of linguistic capacity required for their inclusion among the project team. However, publications with titles given in English but with indication that the main text is in a different language (e.g. "*Agriculture in the Dolisie region, Congo: situation and perspectives on development - the case of a small peripheral town. [French]*") were investigated further to ensure that English language versions were not lost in the process of removing duplicates.
- Relevant special journal issues encountered in the review process were included and all featured articles were considered.
- Other review and synthesis articles were included.

We took an inclusive approach to uncertainty, meaning that publications with titles or abstracts that did not explicitly mention, but could still feasibly meet, all of the above criteria were included for later inspection.

Information extraction and processing

In addition to basic information on each relevant publication (author, year etc.) the following information was collected for each:

- SDG target (and goal) receiving focus. Each target encountered represented a single row in our dataset.
- 'Direction' of the impact identified (beneficial, damaging, mixed, negligible or unknown) (details given in main article).
- The level of confidence associated with each record. Criteria for the three categories is as follows:

    - *Poor* = Based on assumptions or theories only, including predictive models and anticipated impacts. Examples of poor confidence impacts encountered in our review include the work of Chapman et al. (2015)[7], who suggest, but do not demonstrate, that provision of health care can improve community perceptions (and therefore efficacy) of protected areas; the work of Bashaasha et al (2001)[8], whose predictive models suggest that agricultural intensification would reduce forest loss; or the work of Cornet et al. (2018)[9], who anticipated the damaging forest impacts of a new railway link in the UK, which is yet to be built.

    - *Fair* = Based on either: (a) Models that show a correlation between observed progress towards achievement of a goal, or samples that reflect different stages of achievement (e.g. national indicators), and forest change, but with confounding or mediating factors; and/or (b) observations of forest impacts based on qualitative reports or proxy measures (e.g. fuelwood used per household, numbers of people engaged in forest damaging activities etc.). We can illustrate this with the findings of Swinton and Quiroz (2003)[10], who used multiple regression to show that increased levels of education reduced households *reported* likelihood of cutting trees, but that this was only one of several contributing factors, which also included households' physical assets, access to credit, and distance to a paved road, among others.

    - *Good* = Direct observation of forest changes arising from progress towards achievement of a goal, or samples that reflect different stages of achievement (e.g. national indicators). Among the clearest examples of this category are cases where hard infrastructure, such as hydroelectric dams[11], observably results in removal or flooding of forest for its development. A perhaps less obvious illustration is the work of Belay et al. (2015)[12], who showed a direct link between forest regrowth around selected Ethiopian villages (based on remotely sensed images) and the provision of food aid to tackle hunger and malnutrition.

    Literature that provided second-hand (i.e. cited from other sources) records of impacts were earmarked, and confidence levels assigned once all other data collection was complete. This process involved noting the original citation and, where necessary, inspecting the original source for clarification. This was done to avoid duplication of records (i.e. where two or more sources cited the same impact from the same source). Accordingly, if the same source was cited more than once, or if it was already present in our data, it was included only once in our final dataset. This approach, similar to the 'snowball' method, was used only to clarify details of the impact cited in the original citing source, and so we did not consider all records from secondary sources ad infinitum.

- Relevant notes on the impact recorded, including mechanisms by which the impact occurs and any caveats associated with the source material.
- Whether the observed impact was associated with a external intervention.
- Any multi-target impacts recorded or suggested in the paper. This component, however, was not conducted systematically and can only be used for illustrative purposes.

References

1. Inter-Agency and Expert Group in Sustainable Development Goal Indicators. Final list of proposed Sustainable Development Goal indicators. *Rep. Inter-Agency Expert Gr. Sustain. Dev. Goal Indic.* Annex IV (2016).
2. Pullin, A. S. & Stewart, G. B. Guidelines for systematic review in conservation and environmental management. *Conserv. Biol.* **20**, 1647–1656 (2006).
3. Carvalho, G. O. *et al.* Frontier expansion in the Amazon: Balancing development and sustainability. *Environment* **44**, 34–45 (2002).
4. Steward, C. From colonization to 'environmental soy': A case study of environmental and socio-economic valuation in the Amazon soy frontier. *Agric. Human Values* **24**, 107–122 (2007).
5. Fearnside, P. M. Transmigration in Indonesia: Lessons from its environmental and social impacts. *Environ. Manage.* **21**, 553–570 (1997).
6. Baird, I. G. & Shoemaker, B. Unsettling experiences: Internal resettlement and international aid agencies in Laos. *Dev. Change* **38**, 865–888 (2007).
7. Chapman, C. A. *et al.* Providing health care to improve community perceptions of protected areas. *Oryx* **49**, 636–642 (2015).
8. Bashaasha, B., Kraybill, D. S. & Southgate, D. D. Land use impacts of agricultural intensification and fuelwood taxationin Uganda. *Land Econ.* **77**, 241–249 (2001).
9. Cornet, Y., Dudley, G. & Banister, D. High Speed Rail: Implications for carbon emissions and biodiversity. *Case Stud. Transp. Policy* **6**, 376–390 (2018).
10. Swinton, S. M. & Quiroz, R. Is poverty to blame for soil, pasture and forest degradation in Peru's Altiplano? *World Dev.* **31**, 1903–1919 (2003).
11. Fearnside, P. M. Brazil's Samuel Dam: Lessons for hydroelectric development policy and the environment in Amazonia. *Environ. Manage.* **35**, 1–19 (2005).
12. Belay, K. T. *et al.* Spatial Analysis of Land Cover Changes in Eastern Tigray (Ethiopia) from 1965 to 2007: Are There Signs of a Forest Transition? *L. Degrad. Dev.* **26**, 680–689 (2015).