

The impacts of agroforestry interventions on agricultural productivity, ecosystem services, and human well-being in low- and middle-income countries: A systematic review

Sarah E. Castle¹ | Daniel C. Miller¹ | Pablo J. Ordonez² | Kathy Baylis³ | Karl Hughes⁴

¹Department of Natural Resources and Environmental Sciences, University of Illinois at Urbana-Champaign, Urbana, Illinois, USA

²Department of Agricultural and Consumer Economics, University of Illinois at Urbana-Champaign, Urbana, Illinois, USA

³Department of Geography, University of California Santa Barbara, Santa Barbara, California, USA

⁴World Agroforestry Center (ICRAF), Nairobi, Kenya

Correspondence

Sarah E. Castle, Department of Natural Resources and Environmental Sciences, University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA.

Email: saraheb3@illinois.edu

Abstract

Background: Agroforestry, the intentional integration of trees or other woody perennials with crops or livestock in production systems, is being widely promoted as a conservation and development tool to help meet the 2030 UN Sustainable Development Goals. Donors, governments, and nongovernmental organizations have invested significant time and resources into developing and promoting agroforestry policies and programs in low- and middle-income countries (LMICs) worldwide. While a large body of literature on the impacts of agroforestry *practices* in LMICs is available, the social-ecological impacts of agroforestry *interventions* is less well-studied. This knowledge gap on the effectiveness of agroforestry interventions constrains possibilities for evidence-based policy and investment decisions to advance sustainable development objectives.

Objectives: The primary objective of this Campbell systematic review was to synthesize the available evidence on the impacts of agroforestry interventions in LMICs on agricultural productivity, ecosystem services, and human well-being. The secondary objectives were to identify key pathways through which agroforestry interventions lead to various outcomes and how the interventions affect different sub-groups of the population.

Search Methods: This review is based on a previously created evidence and gap map (EGM) of studies evaluating the impacts of agroforestry practices and interventions on agricultural productivity, ecosystem services, and human well-being. We included published and unpublished literature in the English language covering the period between 2000 and October 20, 2017. We searched six academic databases and 19 organization websites to identify potentially relevant studies. The search was conducted for our EGM in mid-2017, and we did not conduct an additional search for this systematic review.

Abbreviations: BACI, before-after, control impact; DID, difference-in-difference; EGM, evidence gap map; FAO, Food and Agriculture Organization; FMNR, farmer-managed natural regeneration; ICRAF, World Agroforestry Center; IPCC, Intergovernmental Panel on Climate Change; IV, instrumental variable; LMICs, low- and middle-income countries; HICs, high-income countries; NGO, nongovernmental organizations; OLS, ordinary least squares; PES, payment for ecosystem services; PICO, population, interventions, comparison type, and outcomes inclusion and exclusion criteria; PRISMA, preferred reporting items for systematic reviews and meta-analyses; PSM, propensity score matching; RCT, randomized controlled trials; RDD, regression discontinuity design; SDGs, Sustainable Development Goals; SM, systematic map; SR, systematic review.

Selection Criteria: We included randomized control trials (RCTs) and quasi-experimental studies assessing the effect of an agroforestry intervention on at least one outcome measure of agricultural productivity, ecosystem services, or human well-being for farmers and their farmland in LMICs. Agroforestry interventions include any program or policy designed to promote and support the adoption or maintenance of agroforestry practices, which include trees on farms, silvopasture, shade-grown crops, and homegardens with trees, among others. Moreover, the studies needed to include a nonagroforestry comparator, such as conventional agriculture or forestry systems or a before-after comparison.

Data Collection and Analysis: We used a standardized data extraction spreadsheet to extract details about each included study. We also used a standardized form to assess risk of bias for each of the included studies in this SR. Meta-analysis techniques were used to combine and synthesize effect size estimates for the outcomes measures that had sufficient data. We used a random effects models for the meta-analyses and use Hedge's g (difference in means divided by the pooled standard deviation) to report effect size estimates. The outcomes without enough evidence for meta-analysis were discussed narratively.

Main Results: We identified 11 studies across nine countries, all of which used quasi-experimental methods. Overall, the quality of the evidence base was assessed as being low. Studies were rated as having high or critical risk of bias if they failed to convincingly address more than one of the main potential sources of bias, namely selection bias, group equivalence, and spillover effects. Given the low number of studies and the high risk of bias of the evidence base, the results of this SR are limited and should be considered a baseline for future work. The results of the meta-analysis for impacts on yields indicated that agroforestry interventions overall may lead to a large, positive impact on yield (Hedge's $g = 1.16 [-0.35, 2.67]$ ($p = .13$)), though there was high heterogeneity in the results ($I^2 = 98.99\%$, $\tau^2 = 2.94$, $Q(df = 4) = 370.7$). There were positive yield impacts for soil fertility replenishment practices, including incorporating trees in agricultural fields and improved fallow practices in fields where there are severe soil fertility issues. In other cases, incorporating trees into the production system reduced productivity and took land out of production for conservation benefits. These systems generally used an incentive provision scheme to economically offset the reductions in yields. The result of the meta-analysis on income suggests that agroforestry interventions overall may lead to a small, positive impact on income (Hedge's $g = 0.12 [-0.06, 0.30]$ ($p = .20$)), with moderately high heterogeneity in the results ($I^2 = 75.29\%$, $\tau^2 = 0.04$, $Q(df = 6) = 19.16$). In cases where improvement yields were reported, there were generally attendant improvements in income. In the cases where payments were provided to offset the potential loss in yields, incomes also generally improved, though there were mixed results for the certification programs and the tenure security permitting scheme. One program, which study authors suggested may have been poorly targeted, had negative yield impacts. There was not enough comparable evidence to quantitatively synthesize the impacts of agroforestry interventions on nutrition and food security outcomes, though the results indicted positive or neutral impacts on dietary diversity and food

intake were likely. Surprisingly, there was little evidence on the impacts of agroforestry interventions on environmental outcomes, and there was no consistency of environmental indicator variables used. However, what has been studied indicates that the environmental benefits are being achieved to at least some extent, consistent with the broader literature on agroforestry practices. The evidence base was insufficient to evaluate the interaction between environmental and social impacts. Several studies explicitly considered variable impacts across different population sub-groups, including differential impacts on small-holders versus large-holders, on woman-headed households versus male-headed households, and on richer groups versus poorer groups. Small-holder farmers typically experienced the most positive effect sizes due to the agroforestry interventions. Women and poorer groups had mixed outcomes relative to men and richer households, highlighting the importance of considering these groups in intervention design.

Authors' Conclusions: There is limited evidence of the impacts of agroforestry interventions, restricting our ability to draw conclusions on the effect sizes of different intervention types. The existing evidence forms a baseline for future research and highlights the importance of considering equity and socio-economic factors in determining suitable intervention design. Some key implications for practice and policy include investing in programs that include pilot programs, funding for project evaluation, and that address key equity issues, such as targeting to smallholders, women, poor, and marginalized groups. Funding should also be given to implementing RCTs and more rigorous quasi-experimental impact evaluations of agroforestry interventions over longer time-periods to collect robust evidence of the effectiveness of various schemes promoting agroforestry practices.

1 | PLAIN LANGUAGE SUMMARY

1.1 | Limited evidence on agroforestry interventions shows positive impacts on agricultural yield and income

Agroforestry interventions may lead to a large, positive impact on yield, though there is high variations in findings. Agroforestry interventions may also lead to a small, positive impact on income. There is insufficient evidence on nutrition, food security and environmental outcomes. Equity concerns of agroforestry interventions appeared in many of the studies, with mixed results, indicating that additional consideration of equity in agroforestry interventions is needed.

1.2 | What is this review about?

Agroforestry is defined as the integration of trees and woody shrubs in crop and livestock production systems. It is widely promoted as a conservation and development tool to sequester carbon, improve soil fertility, and conserve biodiversity on agricultural lands while generating

economic benefits for farmers. Agroforestry is promoted through a range of interventions, including farmer capacity development, provision of tree germplasm, and financial or tenure security provision.

This review examines the evidence on the impacts of any type of agroforestry intervention in low- and middle-income countries (LMICs) on three broad outcomes: agricultural productivity, ecosystem services, and human well-being.

What is the aim of this review?

This Campbell systematic review synthesizes the evidence from 11 quasi-experimental impact evaluations of the impacts of agroforestry interventions in LMICs on agricultural productivity, ecosystem services, and human well-being. The review also identifies key pathways through which agroforestry interventions led to these outcomes, and how the interventions affected different sub-groups of the population.

1.3 | What studies are included?

This review includes studies that evaluate the effect of agroforestry interventions on agricultural productivity, ecosystem services, and human well-being in LMICs. All the studies had methodological weaknesses and high risk of bias. No evaluations using randomised controlled trials for agroforestry interventions were identified.

The review summarises the findings from 11 impact evaluations covering 15 programmes in Brazil, China, Colombia, Indonesia, Kenya, Malawi, Mozambique, Nicaragua, and Zambia.

Eight of the studies evaluated farmer capacity development programmes, which provided extension services and technical training, and five incorporated tree germplasm provision. Three studies evaluated incentive programmes, including payments for ecosystem services (PES) and certification schemes. One study evaluated a policy change, one included a component of market linkage facilitation, and one included community-level campaigning.

The practices that were promoted by the interventions included improved fallow systems, incorporating trees in crop fields, silvopasture, coffee agroforestry, and agrosilvopastoral systems.

1.4 | What are the main findings of this review?

There is a large, positive overall effect of agroforestry interventions on agricultural yields, although there is large variation in the results. The largest positive impacts of agroforestry on yields are associated with less fertile lands, and negative yield impacts are associated with highly productive lands.

There is a very small, positive overall effect of agroforestry interventions on income. Increased or neutral income effects are associated with either increased yields providing additional income, or incentive payments offsetting the costs associated with decreased yields.

Few impact evaluations considered the impacts of agroforestry interventions on nutrition and food security. Qualitative assessment suggests that agroforestry interventions may lead to positive or neutral nutrition and dietary diversity outcomes and may lead to positive food security outcomes.

Few studies considered the impacts of agroforestry interventions on ecosystem services. However, the effects of agroforestry practices on ecosystem services are well-documented in the broader agroforestry literature.

In areas with limited soil fertility, agroforestry interventions provided technical support through extension and training programmes, and in some cases provided access to tree germplasm, to support farmers to adopt agroforestry practices intended to increase yields and incomes. In higher productivity areas, agroforestry interventions provided incentives—such as PES, certification schemes, and tenure security—to adopt agroforestry practices intended for conservation that may reduce overall yields.

1.5 | What do the findings of the review mean?

The existing evidence suggests that there may be positive impacts on agricultural yields and incomes as well as food security and ecosystem services, but appropriate intervention design is dependent on local biophysical and socio-economic characteristics.

1.6 | How up-to-date is this review?

The review authors searched for studies up to October 2017.

2 | BACKGROUND

2.1 | The problem, condition or issue

Agroforestry is widely practiced and promoted across the LMICs of Africa, Asia, and Latin America. Agroforestry practices, ranging from the intercropping of trees within annual crop fields to trees intercropped with plantation species and cultivation of forest gardens, are estimated to take place on nearly 50% of agricultural land in developing country regions (Zomer et al., 2014). Defined simply as “agriculture with trees” or more comprehensively as “the practice and science of the interface and interactions between agriculture and forestry, involving farmers, livestock, trees and forests at multiple scales” (World Agroforestry, 2017), agroforestry comprises an increasingly important strategy to increase food production while advancing other social and environmental objectives.

Proponents argue that agroforestry can provide basic subsistence, natural insurance, and a means to generate income and build assets for many rural households in LMICs (Garrity et al., 2010; Miller, Muñoz-Mora, et al., 2017). Agroforestry can also generate environmental benefits, including carbon sequestration, biodiversity conservation, clean water and improved water infiltration, erosion control and soil fertility, while enhancing resilience of agricultural lands in the face of climate-related stresses (Blaser et al., 2018; FAO, 2013; Garrity et al., 2010; Jose, 2009; Kalaba et al., 2010; Mbow et al., 2014a). In addition, studies also suggest that agroforestry can increase agricultural productivity (Amadu et al., 2020; Sileshi et al., 2008; Waldron et al., 2015) and improve food and nutrition security (Rosenstock et al., 2019; Vira et al., 2015).

Given these benefits, agroforestry has been widely promoted in LMICs. It is expected to play a key role in delivering the UN Sustainable Development Goals (SDGs) (Nations, 2015; Waldron et al., 2017; World Agroforestry, 2017). Government extension agencies, nongovernmental organizations (NGOs), and a range of donor agencies have long provided support to agroforestry systems and practices. Since the 1992 UN Earth Summit in Rio, international aid donors have invested more than US \$10 billion in agroforestry projects (AidData, 2017; activity code: 31220.07) in LMICs (Tierney et al., 2011). The largest donor, the World Bank, continues to emphasize agroforestry in its policy documents, including major

commitments to ensure its agricultural investments are “climate smart” by 2020 (World Bank, 2016). High-level policy documents in many LMICs now explicitly call for the integration of trees into farming systems (e.g., national policies of Government of India (2014), Republic of Kenya (2014), and Government of Malawi (2011)) and there is growing interest in promoting agroforestry as part of sustainable intensification initiatives that reconcile agricultural production with the provision of other important ecosystem services (FAO, 2013; Pretty, 2018).

A large body of literature on the adoption (Mercer & Pattanayak, 2003; Mercer, 2004; Pattanayak et al., 2003) and impacts (Miller et al., 2020) of agroforestry practices in LMICs is now available. However, systematic understanding of the social-ecological impacts of agroforestry *interventions* remains missing. A critical gap exists in knowledge of the on-the-ground effectiveness of interventions promoting the adoption of agroforestry practices in advancing sustainable development priorities. The lack of such knowledge, in turn, hampers the ability of decision-makers to effectively allocate resources relating to agroforestry research, policy, and practice. This systematic review (SR) addresses this need for evidence synthesis, focusing on impact evaluations that assess the effects of agroforestry interventions on agricultural productivity, ecosystem services, and human well-being.

2.2 | The intervention

Agroforestry is promoted and supported in a variety of ways. Previous work (Miller et al., 2020) has identified six main agroforestry intervention categories:

- *Farmer capacity development* through training, extension, the provision of other advisory services and technical information, demonstration sites, participatory trials, and other modes of action learning.
- *Incentive provision* through direct payments to farmers for planting and caring for trees on their farms and the receipt of premiums for particular agricultural commodities, e.g., for shade grown coffee.
- *Enhancing access to tree germplasm* through the direct provision of tree seedlings/seeds and linking farmers to and/or strengthening the capacity of tree germplasm suppliers.
- *Community-level campaigning and advocacy* encouraging large numbers of community members to plant trees on their farms and/or pursue specific agroforestry practices.
- *Market linkage facilitation* for a greater and/or more favorable integration of smallholders into tree-product value chains.
- *Policy and institutional change* for a more enabling environment that promotes the uptake of agroforestry and/or enables its potential benefits to be better realized.

Agroforestry interventions typically encourage farmers to take up several complementary practices (e.g., planting of longer-term tree species together with short-term shrubs along field contours) to meet multiple social-ecological objectives (Waldron et al., 2017). The

establishment of trees incorporated into crop fields or pasture, trees integrated with plantation crops, and improved or rotational fallow are other common examples of promoted practices, which may include the provision of training and material support in setting up of tree nurseries and grafting stock. Strengthening the integration of smallholders into tree-product value chains through, for example, addressing production constraints or promoting more favorable contractual arrangements with buyers, is also increasingly popular (Degrande et al., 2014).

2.3 | How the intervention might work

A simplified and generic theory of change that may underlie an agroforestry intervention (either explicitly or implicitly) is presented in Figure 1. The first required step is successful mobilization and engagement of farmers or landholders - those that would potentially adopt new or expanded agroforestry practices on their land. The second step represents a given interventions, such as farmer capacity development or facilitating access to appropriate tree germplasm. At least the first and, in many cases, both are required for significant and appropriate adoption of the promoted agroforestry practices and/or tree germplasm. Following such adoption, several intermediary outcomes are then expected. For example, farmers may see improved soil health and other ecosystem services, such as water infiltration, that then increase crop productivity or reduce production costs and, therefore, increase returns. Some participants in the intervention may find that increased use and availability of tree/shrub fodder leads to increases in milk and other livestock production and returns. Selling other agroforestry products such as timber, firewood, and fruit, is also expected to increase and diversify income and food sources (Mbow et al., 2014b; Sharma et al., 2016; Waldron et al., 2017). These changes may have differential effects depending on gender. Together, these intermediate outcomes are expected to interact together to bolster household resilience to shocks, as well as overall household income food and nutritional security. These positive benefits—and the broader context in which this stylized theory of change is embedded—will then affect further household investment in agroforestry.

2.3.1 | Potential tradeoffs

Our theory of change diagram presents positive pathways linking agroforestry interventions, adoption, and beneficial impacts. However, agroforestry may also include potentially negative tradeoffs, such as a reduction in area of crop production and negative tree-crop interactions. Though the evidence is mixed (e.g., Blaser et al., 2018), a reduction in productivity may accompany agroforestry practices. Therefore, some interventions promoting new or expanded agroforestry practices may require a mode of compensation for yield losses. Such compensation may come in the form of PES or certification programs that yield higher prices for the crops

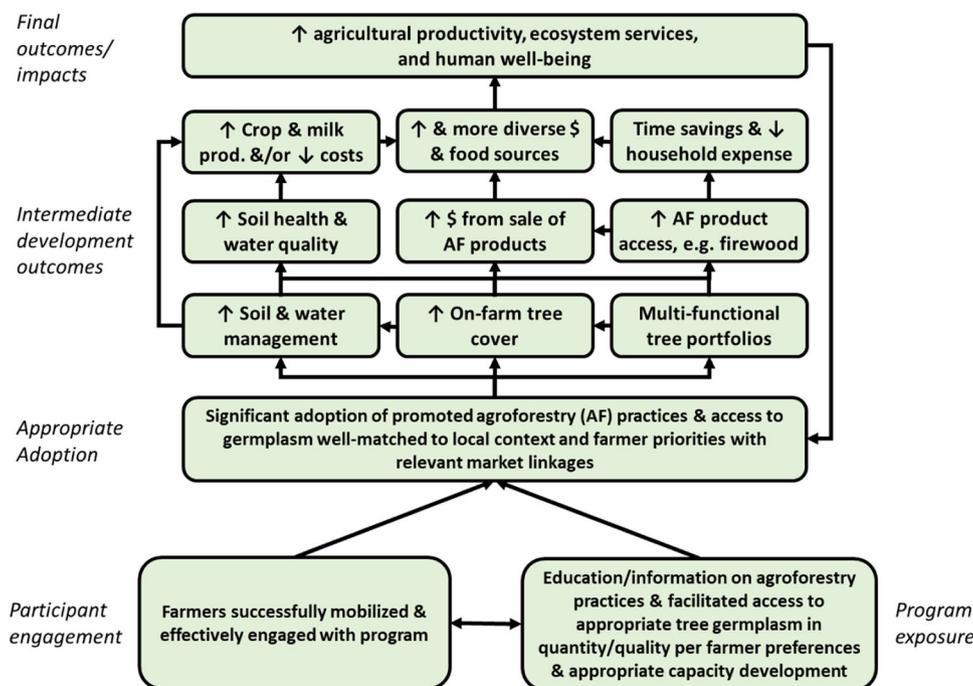


FIGURE 1 Illustrative theory of change for an AF intervention. Figure presented in Miller et al. (2020) and used with permission here. AF, agroforestry

the farmers produce. These types of interventions may help balance the tradeoff between environmental benefits, like biodiversity conservation, soil and water quality, and carbon sequestration, with agricultural yield and economic ones. Furthermore, while there may be short-term tradeoffs with reductions in yields, this may not stay true in the long term as productive tree crops reach maturity or as soil fertility increases (Garrity et al., 2010; Nair, 1993; Pandey, 2007). The role of climate change further affects potential tradeoffs in agroforestry systems. Such systems may provide, for example, climate change resilience, which may result in productivity advantages during difficult years with extreme weather events.

2.4 | Why it is important to do this review

Agroforestry systems and practices are found across LMICs and are viewed as increasingly important for boosting food security, addressing environmental degradation, and contributing to a range of other development policy objectives (Garrity et al., 2010; Waldron et al., 2017). However, financing and implementation of agroforestry and other nonmainstream agricultural approaches remains limited in many contexts (DeLonge et al., 2016; Horlings & Marsden, 2011; IPES-Food, 2016). Instead, high-input, mechanized approaches to agriculture predominate. Over the past half century, these approaches have become conventional, leading to major increases in yields and helping to feed much of the world's population (IAASTD, 2009; Pretty & Bharucha, 2014; The Government Office for Science, 2011). However, these benefits

have brought with them sometimes steep social and environmental costs, including biodiversity loss, climate change, land degradation, water pollution, and negative effects on human health (Brawn, 2017; Horrigan et al., 2002; IAASTD, 2009; Intergovernmental Panel on Climate Change, 2015; Maxwell et al., 2016; Pretty & Bharucha, 2014).

Farmers, consumers, and policymakers increasingly recognize these environmental and health costs and seek viable alternatives that can simultaneously address food security concerns while delivering other social and environmental benefits. Agroforestry represents one such alternative, but there is an important need to systematically identify what kinds of interventions and practices have worked to deliver these benefits and understand potential trade-offs involved. Evidence on the effectiveness of agroforestry interventions is therefore needed to inform broader debates and investment decisions relating to sustainable agricultural intensification. We expect that this SR will present a vital resource to inform such discussions, including on expanded measures that account for the multiple values of agroforestry and other agricultural systems (Sukhdev, 2018; Waldron et al., 2017).

This SR uses evidence compiled in a recently published EGM on the impacts of agroforestry practices and interventions on agricultural productivity, ecosystem services, and human well-being (Miller et al., 2020). In this mapping exercise, we identified 395 studies on the impacts of agroforestry practices and interventions, including eight impact evaluation of agroforestry interventions and 11 SRs. An extended search identified an additional three impact evaluations included in this SR.

All the SRs we found on agroforestry study the impact of agroforestry practices, without considering interventions promoting and supporting the adoption of agroforestry leading to social-environmental outcomes. These SRs include Reed et al. (2017) on the impact of trees on food production and livelihoods in the tropics; meta-analyses of agricultural yields with and without trees in West Africa (Bayala et al., 2012; Sileshi et al., 2008); a global meta-analysis of agroforestry impacts on pasture yields (Rivest et al., 2013); global meta-analyses of the carbon sequestration potential of agroforestry (Kim et al., 2016) and on soil carbon storage (Corbeels et al., 2018); biodiversity functions of agroforestry in the tropics (Jezeer et al., 2017; Norgrove & Beck, 2016) and globally (De Beenhouwer et al., 2013); a meta-analysis on the use of trees in agriculture on infiltration capacity (Ilstedt et al., 2007); and a global meta-analysis of the impacts of agroforestry on pest, disease, and weed control (Pumariño et al., 2015).

As detailed below, the current SR includes all LMICs, not just tropical ones, and both direct and indirect effects of agroforestry interventions on a range of outcomes, including multi-dimensional human well-being. We are aware of no SR that summarizes empirical studies on the causal effects of agroforestry interventions in LMICs, particularly outside the context of tightly controlled, research station-based experimental trials.

There are two primary audiences for this SR. First, we expect that researchers on agroforestry and broader sustainability issues will use the results to inform further investigations on these topics, including new empirical research. Results should be of wide interest to researchers in a range of institutions, from CGIAR centers to universities. The second main anticipated audience is decision-makers for whom agroforestry is already or potentially of interest. This includes relevant ministries and programs in governments and donor agencies, as well as NGO and other advocacy and implementing organization staff.

3 | OBJECTIVES

The overall aim of this SR is to identify and synthesize existing evidence on the effects of interventions that promote the adoption and use of agroforestry practices on agricultural productivity, ecosystem services and human well-being in LMICs.

In this SR, we address the following three research questions:

- 1) What effects do agroforestry interventions have on agricultural productivity, ecosystem services, and human well-being outcomes?
- 2) What are the effects of agroforestry interventions on different study population sub-groups?
- 3) What are the pathways through which agroforestry interventions generate impacts?

4 | METHODS

4.1 | Criteria for considering studies for this review

The included studies in this SR were identified based on results from our evidence gap map (EGM) of the impacts of agroforestry on agricultural productivity, ecosystem services, and human well-being in LMICs (Miller et al., 2020). We based our EGM, and thus this review, on a previously published protocol (Miller, Ordonez, et al., 2017). Here, we summarize the methods used in that EGM and then present methods used specifically to carry out this SR. We note that the EGM included studies that evaluated the impacts of agroforestry practices, but our SR only considers those studies that evaluated the impacts of agroforestry interventions. The selection criteria for studies in this review are summarized in Table 1 and discussed in detail below.

4.1.1 | Types of studies

This SR includes quantitative impact evaluations using experimental or quasi-experimental designs. Experimental designs use random assignment to treatment and control groups, such as a randomized controlled trial (RCT). Quasi-experimental designs use rigorous statistical methods to adjust for nonrandom assignment between treatment and control groups to make causal inferences. We include the following types of quantitative impact evaluation studies (Snilstveit et al., 2019):

- Studies where participants are randomly assigned to treatment and comparison group (experimental study designs);
- Studies where assignment to treatment and comparison groups is based on other known allocation rules, including a threshold on a continuous variable (regression discontinuity designs) or exogenous geographical variation in the treatment allocation (natural experiments);
- Studies with nonrandom assignment to treatment and comparison group that include pre-and post-test measures of the outcome variables of interest to ensure equity between groups on the baseline measure, and that use appropriate methods to control for selection bias and confounding. Such methods include statistical matching (for example, propensity score matching (PSM), or covariate matching), regression adjustment (for example, difference-in-differences, fixed effects regression, single difference regression analysis, instrumental variables (IVs), endogenous switching regression, and “Heckman” selection models);
- Studies with nonrandom assignment to treatment and comparison group that include post-test measures of the outcome variables of interest only and use appropriate methods to control for selection bias and confounding, as above.

TABLE 1 Elements of the agroforestry systematic review

Population	Intervention	Comparators	Outcomes
Farmers and/or farmland in low-and middle-income countries	Implementation of one or more of the defined agroforestry interventions	Control group of nonparticipants; or, before-after time-series comparison	Positive, negative, or neutral effects on agricultural productivity, ecosystem services, or human well-being

Ideally, studies would have included both baseline and post-intervention data. However, given the small number of studies meeting this criterion, we include studies with only post-intervention outcome data as long as they use some method to control for selection bias and potential confounding factors.

We excluded theoretical or modeling studies (unless they include a relevant empirical example with design that meets inclusion criteria), editorials and commentaries, and field trials that were not part of a specific intervention.

4.1.2 | Types of participants

The population of interest was farms and those that live and farm on them in LMICs using a system that falls within the definition of agroforestry.

4.1.3 | Types of interventions

From a policy perspective, it is especially useful to know what kinds of interventions might most effectively promote agroforestry practices to yield desired social-ecological outcomes. This SR focuses on the types of interventions summarized in Table 2.

The promotion of agroforestry includes a wide range of specific practices that fall under what is generally considered as agroforestry. Here, we consider “agroforestry” to be defined as “a collective name for land-use systems and technologies where woody perennials (trees, shrubs, palms, bamboos, etc.) are deliberately used on the same land-management units as agricultural crops and/or animals, in some form of spatial arrangement or temporal sequence” (Nair, 1993). To capture the wide diversity of practices that might fall under this definition and present them in a coherent way, we subdivided agroforestry into the practice types listed in Table 3. This set of practice types is based on the classification system proposed by Nair (1985, 1993) and updated by Sinclair (1999), Torquebiau (2000), and Atangana et al. (2014).

To identify the effect of an intervention or practice, a study needs to include both adopters or those exposed to an agroforestry intervention and comparators. A comparator is defined as a farm or household that does not adopt a given practice identified in Table 3, or is not exposed to a specific agroforestry intervention. Specifically, eligible comparisons included a land or household where agroforestry was not practiced but another land use was in place (e.g., agriculture, primary forest, or secondary forest/forest plantation).

For observational studies, a farm or household before adopting a given agroforestry practice was also an eligible comparator.

4.1.4 | Types of outcome measures

This SR focuses on three broad outcome categories: (1) agricultural productivity, (2) ecosystem services, and (3) human well-being. Studies that focused exclusively on the adoption of a particular agroforestry technique or species without reference to impact were excluded. We did not specify a minimum or maximum duration of follow-up for study inclusion. All types of agricultural production settings in LMICs were considered relevant.

Importantly, we excluded studies that evaluated the impact of agroforestry interventions on adoption only without estimating the impacts on any measure of our three broad outcome categories. We believe these outcomes are most of interest to policymakers considering agroforestry interventions. Additionally, while the adoption of agroforestry due to an intervention may be an important indicator for the longer-term effects of the intervention given the larger body of evidence for the impacts of agroforestry practices, the realized outcomes are highly uncertain. However, agroforestry practices are highly diverse, there is high variability in the long-term outcomes, and the impacts are context-specific (Coe et al., 2014). One practice may lead to very different outcomes in different contexts (Friggens et al., 2020; Martin et al., 2020). Furthermore, adoption alone, especially when only considering early adoption, says very little about the effectiveness of an intervention. Agroforestry impacts depend on tree survival, farmer commitment to maintaining the practice, and the biophysical features of the land. Therefore, this review focuses on the studies measuring outcomes beyond adoption. Future research may examine adoption-only studies.

4.2 | Search methods for identification of studies

The search and screening process was conducted for the agroforestry EGM by Miller et al. (2020), based on a previously published protocol for the EGM (Miller, Ordonez, et al., 2017). A confirmatory search for this SR, led us to identify an error in the initial EGM search string. To address this issue, we conducted an additional search with the corrected search string. These changes are documented in Appendix A.

TABLE 2 Classification of interventions to promote agroforestry

Intervention type	Description and examples
Farmer capacity development	Efforts focus on enhancing farmer knowledge and/or skills relevant to agroforestry practice, for example, setting up and managing tree nurseries; tree planting and management techniques; and seed collection and propagation. Such interventions can involve the provision of training, extension and other advisory services, and specific technical information, as well as the setting up of demonstration sites, running of participatory trials and other modes of participatory action learning
Enhancing access to tree germplasm	Efforts to facilitate farmer access to quality and desired tree/shrub seedlings/seeds required to pursue prioritized agroforestry practices. Such interventions often entail the direct provision of seedlings/seeds to farmers but can also involve linking farmers to relevant suppliers and/or enhancing the ability of existing or new suppliers to supply participating farmers with quality and desired tree germplasm
Community-level campaigning and advocacy	Interventions of this type can also involve the provision of information about the benefits of trees and agroforestry and/or the provision tree seedlings/seeds but is distinct from the first two types. The main objective is to motivate, including through social pressure, community members to plant trees on their farms and/or pursue specific agroforestry practices. Campaigning and advocacy may be done through radio and/or community meetings, speeches, and drama and may involve a mass community effort to plant trees, for example, on a specific day of the year
Incentive provision	Interventions of this type seek to motivate farmers to plant trees and practice agroforestry through the provision of incentives. Examples include paying farmers for planting and caring for trees on their farms in exchange for desired ecosystem services (e.g., carbon sequestration) and buyers offering premiums to farmers for agricultural commodities produced under certain conditions (e.g., via certification schemes for products such as shade grown organic coffee)
Market linkage facilitation	Interventions of this type focus on efforts to enhance potential returns from agroforestry to encourage adoption. This could be through linking producers to and/or brokering new and/or improving existing contractual arrangements with buyers. Other examples include the collective marketing of agroforestry products and/or interventions to stimulate demand for a given agroforestry product, for example, Baobab fruit
Institutional and policy change	Interventions of this type involve reforming and/or putting in place new policies, laws, regulations, and institutions more broadly to facilitate greater uptake of and benefits from agroforestry. Such efforts are designed to address existing policy and institutional constraints such as, for example, prevailing forestry regulations—designed for forest management areas—that may frustrate smallholder efforts to grow particular high-return tree species or insecure land tenure that may similarly deter long-term investments in tree planting

Studies from the year 2000 to October 20, 2017 were included in the EGM search. We began the study period in 2000 given that year marked the start of the Millennium Development Goals, which presaged the current SDGs, and that agroforestry had gained significant momentum by then in the wake of the 1992 UN Earth Summit in Rio. The search was carried out in October 2017, and we included only English language studies due to resource constraints. We included studies irrespective of their publication status. We searched the citations and references of the included SR studies, but we did not find evidence of any additional papers that meet our inclusion criteria within our SR time frame from this additional search.

The search details, including a list of the databases and organizational websites and the example full search string, are provided in Appendix A.

4.3 | Data collection and analysis

The online literature review and reference management software, EPPI-Reviewer 4, was used to upload relevant titles and abstracts for

candidate studies identified through the search strategy for our EGM. The EGM specifically marked whether a study considered an agroforestry intervention (versus only a practice) and if it used experimental or quasi-experimental methods. These impact evaluations of interventions comprise the evidence base for this SR. The information for each study included in this SR was extracted into a data extraction matrix in Excel. Our data extraction matrix was adapted from the data extraction matrix used in Snilstveit et al. (2019). The data we extracted included bibliographic information, study design, context, intervention information, process and implementation, cost, external validity, outcome information, and outcome data to be used in the meta-analyses. The data extraction matrix is presented in Supporting Information Appendix 2.

4.3.1 | Selection of studies

For our EGM, we imported the records from academic databases into our data management software (EPPI-Reviewer 4), and we used the built-in tool to aid in removing duplicates. The grey literature was

TABLE 3 Classification of agroforestry systems and specific practices

Agroforestry system	Specific Practices	Definition
Agrisilvicultural (crops and trees)	Improved or rotational fallow	Land resting system using trees and shrubs to replenish soil fertility, sometimes in rotation with crops as in traditional shifting cultivation
	Multipurpose trees on parklands or lots (mixed trees and crops)	Scattered trees in parklands (landscapes derived from agricultural activities) or other land area or in systematic patterns on bunds, terraces or plot/field boundaries
	Mixture of plantation crops	Combination of plantation crops in an intercropping system in alternate arrangement, including use of shade trees for cash crops
	Tree gardens	Cultivation of a mixture of several fruit and other useful trees, sometimes with the inclusion of annual crops. This arrangement is sometimes referred to as homegardens
	Alley cropping	Planting rows of trees with a companion crop grown in the alleyways between the rows
	Shelterbelts	Extended windbreak of living trees and shrubs established and maintained to protect farmlands (beyond a single farm)
	Silvopastoral (pasture/animals and trees)	Multipurpose fodder trees or shrubs around farmlands (protein bank)
Agrosilvopastoral (crops, pasture/animals and trees)	Living fences and shelterbelts	Trees as fences around plots and/or an extended windbreak of living trees and shrubs established and maintained to protect farmlands and provide fodder
	Integrated production of animal/dairy and wood products	Production of animal/dairy and wood products within the same land area
	Trees/shrubs on pasture	Trees scattered irregularly or arranged according to some systematic pattern
	Integrated production of animals (meat and dairy), crops, and wood/fuelwood	Production of crops, animal/dairy and wood products within the same land area, including around homesteads
	Woody hedgerows for browse, green manure, soil conservation	Multipurpose woody hedgerows for browse, mulch, green manure, soil conservation, and so forth
Agroforestry including insects/fish	Wooded pasture products	Land covered with grasses and other herbaceous species, and with woody species
	Entomoforestry	The combination of trees and insects (e.g., bees for honey and trees)
	Aqua-silvo-fishery	Trees lining fish ponds, tree leaves being used as "forage" for fish

imported into and managed in Microsoft Excel due to reference format incompatibility with EPPI-Reviewer 4. We screened the records at the title and abstract level, excluding studies which did not meet our criteria for study country, publication year, study type, and relevant agroforestry practice or intervention.

The review process consisted of 14 reviewers. All the reviewers were trained by the project leads (Miller and Baylis) and research coordinators (Ordoñez and Castle). We first reviewed search results at the level of title and abstract to determine inclusion or exclusion. The title and abstract stage of the review process included 11 reviewers. To ensure inter-rater reliability, each reviewer was given two samples of 30 studies for classification (60 studies in total). Results from one of the lead researchers was used as the standard for classification and a kappa statistic was used as a measure of agreement between reviewers (Cohen, 1960). This statistic was calculated for each reviewer against the standard classification. At least a 70% agreement was required for all reviewers. If the initial sample did not yield the required agreement, reviewers would discuss their responses with a project lead and retake the test until the required agreement level was reached. Once the review process started, if a reviewer was unsure about the inclusion of a given study, the reviewer had the option to mark it for a second opinion. The research leads and coordinators made inclusion decisions in such cases. In addition, these same reviewers performed a second title and abstract screening of all studies marked for inclusion, at which point some additional studies were excluded that were found not to meet the inclusion criteria. The review team met weekly during the entire screening process, during which time reviewers discussed questions and agreed on coding decisions for studies that a given reviewer was uncertain whether to include. The research coordinators continuously verified that a subset of screened studies (~10% of studies coded by each reviewer) were coded correctly to ensure inter-reviewer reliability was maintained.

The full team of reviewers then screened remaining studies at the full text level. At this stage, reviewers also had the option to mark studies for second opinion, with the lead researchers making final determinations as in the previous stage. Throughout this and the full-text stage, the research team met regularly to discuss any issues or inconsistencies, and the lead researchers/research coordinators did spot-checks. Approximately 70% of the studies were double-screened at the full-text level by one reviewer and one research coordinator. Additionally, all the studies included were subject to data extraction checks by the research coordinators. For approximately 70% of the included studies, the full data extraction performed by a reviewer was verified by a research coordinator, while the remainder were spot-checked by a research coordinator. Any disagreements on coding or data extraction were discussed between the reviewer and coordinator, and if an agreement could not be reached, the other research coordinator made the final decision.

During the full data extraction process for the EGM, we marked whether a study (1) considered an agroforestry intervention, instead of only an agroforestry practice (40 out of the 384 primary studies), and (2) whether the study used an experimental or quasi-experimental

study design, instead of observational only (10 out of the 384 primary studies). Of these two criteria, eight studies met both criteria to be included as an experimental or quasi-experimental impact evaluation of an agroforestry intervention. These eight studies comprise our initial evidence base for this SR. Our additional search for this SR yielded an additional three included studies.

4.3.2 | Data extraction and management

We conducted data extraction that expanded on the related EGM for the included studies in this SR. We used a standardized data extraction form, including a codebook describing the scope of each question on which we sought data, to compile descriptive data from all studies meeting our inclusion criteria (Supporting Information Appendix 2). We extracted the following types of information:

- Bibliographic information.
- Study design and basic information, including information on the intervention type, funder, implementing agency, intervention objectives, location, details on the program, target groups, number of participants, duration, follow up, practices promoted, sample size of treatment and control group, comparators, and equity focus groups.
- Process and implementation, including information on program uptake and adherence, implementation fidelity and service delivery quality, and other process factors.
- Cost.
- External validity measures, including length of study, conditions of trial, independence of evaluation, conflicts of interest, and use of theory.
- Outcome measures, including types of outcomes evaluated, indicator variables, equity groups examined, sample size, effect size data, and standard deviation.
- Mechanisms, including any stated mechanisms linking the intervention to the outcome.

One person (Castle) undertook this descriptive data extraction and a senior reviewer checked it for agreement. Two reviewers (Castle and Ordoñez) undertook the effect size data extraction to check for consistency.

4.3.3 | Assessment of risk of bias in included studies

Following Snilstveit et al. (2019), we undertook risk of bias assessments of each of the included impact evaluations using criteria as suggested by an adapted version of the Cochrane Risk of Bias Tool. We assessed the risk of bias by coding “Yes,” “No,” and “Unclear” for each of the following criteria:

1. Mechanism of assignment: was the allocation or identification mechanism able to control for selection bias?

2. Group equivalence: was the method of analysis executed adequately to ensure comparability of groups throughout the study and prevent confounding?
3. Performance bias: was the process of being observed free from motivation bias?
4. Spill-overs, cross-overs and contamination: was the study adequately protected against spill-overs, cross-overs and contamination?
5. Selective outcome reporting: was the study free from selective outcome reporting?
6. Selective analysis reporting: was the study free from selective analysis reporting?
7. Other risks of bias: Is the study free from other sources of bias?

Two separate reviewers assessed each study for risk of bias independently. A third reviewer resolved any disagreements. The results of this analysis are reported for each of the assessed criteria for each study.

Again following Snilstveit et al. (2019), we used the results of the risk of bias assessments to produce an overall rating for each study as low, medium, high or critical risk of bias. We used the following decision rules to come to this decision:

- If all questions are answered “yes,” studies are assigned a low risk of bias rating.
- If studies score “yes” for selection, group equivalence and spillovers, but “no” or “unclear” for other domains studies are assigned a medium risk of bias rating. If they score “yes” for two out of three of the categories selection, group equivalence and spillovers, and unclear for another, we assign a medium risk of bias rating.
- If studies score “no” for any one of the following: selection, group equivalence or spillovers they are assigned a high risk of bias rating. For studies unclear on two or more of the three key categories (selection, group equivalence or spillovers) but that attempted matching/matching w. regression, we give a high risk of bias rating.
- If studies score “no” for more than one of the selections, group equivalence or spillover questions the study is assigned a critical risk of bias rating.
- Otherwise, we take an unclear rating as “no.”

4.3.4 | Measures of treatment effect

The statistical evidence in the studies was extracted with the intention of comparing the estimated effects of interventions on outcomes. Two reviewers independently extracted the data from a random sample of studies to ensure consistency and resolve any inconsistencies.

We extracted outcome data from each study and calculated standardized effect sizes, standard errors, and confidence intervals. For continuous outcomes, we calculated standardized mean difference effect sizes using Hedges' g (sample size corrected) standardized mean difference (SMD). To adjust for the small positive bias

resulting from the Hedges' g calculation, we use the following equation to obtain an unbiased version of Hedges' g (Borenstein et al., 2009b; Ellis, 2010; Hedges & Olkin, 1985):

$$g = d \times J,$$

where J is the correction factor

$$J = 1 - \frac{3}{4(n_t + n_c) - 9}.$$

To calculate the effect size, d , we use one of the following formulas, based on the outcome reporting provided by each study (Borenstein et al., 2009b).

For studies reporting mean difference and standard deviations, we calculate the effect size according to the following formula:

$$d = \frac{\bar{X}_t - \bar{X}_c}{SD_{\text{pooled}}^*},$$

where the pooled standard deviation is

$$SD_{\text{pooled}}^* = \sqrt{\frac{(n_t - 1) * SD_t^2 + (n_c - 1) * SD_c^2}{n_t + n_c - 2}}.$$

For studies reporting a regression coefficient, β , between continuous variables, the regression coefficient to standardized mean difference, g using the following equation:

$$g = J \times \frac{\beta}{SD},$$

where the standard deviation, SD , is

$$SD = SE * \sqrt{N} = \frac{\beta}{T.\text{value}} * \sqrt{N},$$

where SE is the standard error, calculated as the regression coefficient divided by the T value for the coefficient, and N is the total sample size in the regression model.

The variance is

$$V_g = J^2 \times \left(\frac{n_t + n_c}{n_t n_c} + \frac{d^2}{2(n_t + n_c)} \right).$$

And the standard error is

$$SE_g = \sqrt{V_g}.$$

For studies reporting a correlation coefficient between two continuous variables, the correlation coefficient, r , serves as the effect size index. We can convert r to standardized mean difference, g using the following equation:

$$g = J \times \left(\frac{2r}{\sqrt{1 - r^2}} \right).$$

And the variance is

$$V_d = \frac{4}{(1 - r^2)(n - 1)}.$$

Hedge's g is typically interpreted using the following approximations, though the meaning of small, medium, and large depends on the context (Cohen, 1977):

- Small effect (cannot be discerned by the naked eye) = 0.2
- Medium effect = 0.5
- Large effect (can be seen by the naked eye) = 0.8

4.3.5 | Unit of analysis issues

Several of our included studies reported dependent effect sizes, but only effect sizes that are statistically independent should be used in meta-analysis. Specifically, three studies reported effect sizes for multiple treatment groups within the same study (Haggart et al., 2017; Pender et al., 2008; Thorlakson & Neufeldt, 2012). Additionally, one study used panel data and presented results for multiple outcomes at multiple time points (Sills & Caviglia-Harris, 2015). To deal with effect size dependence due to multiple comparisons, multiple outcomes, or multiple time points within a study, we followed the methods presented in Borenstein et al. (2009a, 2009c). We compute a summary effect for the different treatment groups that use the same control group reported within a study to use in our meta-analyses. Since each study will be represented by one summary effect size in the meta-analysis, we avoid more weight being given to studies reporting more treatment groups. The variance of the summary effect size incorporates the correlation among the treatment groups.

The summary effect size is computed as the mean of the effect sizes for each treatment group.

$$\bar{Y} = \frac{1}{n} \sum_{i=1}^{i=n} Y_i,$$

where n is the number of treatment groups, and Y_i is the effect size for treatment group $i \in n$.

The summary effect size variance is:

$$V_{\bar{Y}} = \left(\frac{1}{n}\right)^2 \text{var}\left(\sum_{i=1}^{i=n} V_i\right) = \left(\frac{1}{n}\right)^2 \left(\sum_{i=1}^{i=n} V_i + \sum_{i \neq k} r_{ik} \sqrt{V_i} \sqrt{V_k}\right),$$

where V_i is the variance for each treatment group effect, V_k is the variance for every other treatment group effect size, and r_{ik} is the correlation between each treatment group effect sizes ($r_{ik} = 0.5$ when considering multiple treatments using the same control group).

When a study reported on multiple outcomes measures for a single outcome type for our analysis, we selected a single outcome measure that was most similar to the measures used in the other studies included in our meta-analysis, as we describe in Section 4.3.9.

4.3.6 | Dealing with missing data

One of the included studies did not provide sufficient data to calculate effect sizes. We contacted study corresponding author when there was missing or incomplete data for calculating effect sizes, who provided the needed information.

4.3.7 | Assessment of heterogeneity

We used a random-effects model and report the I^2 statistic to assess the percentage of variability in the estimates due to heterogeneity, the τ^2 variation in the observed effects, and Q statistic difference between the observed effects and fixed-effects model estimates (Higgins et al., 2020). The forest plots also provide a graphical visual of the heterogeneity.

4.3.8 | Assessment of reporting biases

We did not explicitly assess reporting biases due to the small number of studies included in our review. If we had included at least 10 studies in a meta-analysis, we would have reported funnel plots with tests for funnel plot asymmetry (Higgins et al., 2020).

4.3.9 | Data synthesis

We synthesized the study results using conventional meta-analysis (Borenstein et al., 2009b) where data permitted, and narrative synthesis for all studies meeting our inclusion criteria. For synthesis purposes, we looked at impacts of agroforestry on agricultural productivity (yield), income, human nutrition and well-being, and environmental outcomes. We attempted to reduce publication bias by searching for and including unpublished studies and grey literature in the review.

We used the “metafor” package in R software to conduct meta-analysis and create forest plots with effect sizes from each included study. We used the random-effects model in the metafor package for the meta-analyses since we expect that the true effect will vary from study to study included in our meta-analyses (Borenstein et al., 2009b; Viechtbauer, 2010). When a study reported effect sizes for multiple similar outcome measures or when it reported effect sizes based on multiple matching techniques, we reported in our meta-analysis the approaches most similar to the other effect size approaches in the other studies. The included studies for the meta-analyses used PSM, endogenous switching regression, and IVs techniques. When the results of more than one matching technique were reported for the studies using PSM, we used the results of PSM with Kernel Matching techniques. When a study presented both PSM results and endogenous switching regression results separately, we use the results from endogenous switching regression in the meta-analysis and discussed the differences narratively. When a study

presented both PSM and IV results separately, we used the PSM results in the meta-analysis. Some studies also used more than one method as part of their analysis, such as IV with endogenous switching regression. We assessed heterogeneity in the effect size results graphically by presenting the effect size distribution on forest plots for the meta-analyses.

We decided to perform meta-analysis when there were three or more studies presenting comparable indicators for a given outcome. For the outcomes that do not have enough data to perform meta-analysis or where outcome measures were not comparable, we provide a narrative discussion of the trends in size and direction of effect sizes from the studies as well as a discussion of the mechanisms suggested to link the intervention with the outcome. The narrative discussion also highlights difficulties in measuring the outcomes and methods used as well as how comparable the evidence is across the included studies.

The risk of bias assessment was used to determine the overall quality of the evidence base. A “low” or “medium” risk of bias assessment would be considered higher quality evidence in terms of controlling sources of bias than a study with “high” or “critical” risk of bias. We did not use the risk of bias in the meta-analyses and did not restrict our analyses based on risk of bias rating since all of the studies we identified were rated as a high or critical risk of bias. We discuss throughout this report the importance of considering the low overall quality of the evidence base when interpreting the results, and we highlight these points in our discussion. Our primary conclusions include the overall limited evidence base on agroforestry interventions, which is in part drawn from our risk of bias assessments. The results are presented in a table and discussed the assessment narratively.

4.3.10 | Subgroup analysis and investigation of heterogeneity

We planned to conduct qualitative investigation of subgroup heterogeneity across multiple dimensions of equity, specifically gender, race/ethnicity, socio-economic status, and literacy/education level. In our subgroup analysis, we only reported on gender and socio-economic status due to a lack of studies reporting on race/ethnicity or literacy/education level.

In our data extraction, we noted any reference to equity in the included studies. Equity focus is defined as the extent to which an intervention or analysis focuses on specific disadvantaged populations. We aimed to identify how and to what extent the included studies considered equity in their approach. We used the PROGRESS framework (O'Neill et al., 2014) to consider potentially disadvantaged groups in the included studies. Key dimensions of equity that we considered were gender, race/ethnicity, socio-economic level, and literacy/educational level. We assessed the extent to which each study addresses equity, by describing any intervention focus on specific social groups, examining equity as an outcome, or reporting on differential impacts across sub-populations.

4.3.11 | Sensitivity analysis

Due to the limited studies in our review and overall high risk of bias present in the included studies, we were unable to conduct any sensitivity analyses to assess the effects of risk of bias on the treatment effect estimates.

4.3.12 | Treatment of qualitative research

We did not include qualitative research; however, we would have included qualitative studies in our review had they met our inclusion criteria (Miller, Ordonez, et al., 2017). We did not identify any qualitative experimental or quasi-experimental impact evaluations of any agroforestry interventions.

5 | RESULTS

5.1 | Description of studies

5.1.1 | Results of the search

For the EGM, the search returned 20,271 records, with 16,535 studies remaining for screening at title and abstract after duplicate removal. After title and abstract screening, 1557 studies had to be reviewed at full text. Of these, 396 met the inclusion criteria for the EGM. The main reasons for exclusion were lack of relevant intervention/practice ($n = 6750$) and type of study ($n = 4898$). The relatively low number of remaining studies ($n = 963$) were excluded for other reasons, for example, year, country, comparator, or outcome. Only 11 of the 4017 studies identified from the grey literature sources were included in the final EGM. An additional search was conducted in October 2020 with a correction to the initial EGM search string, which returned an additional 6874 studies published through October 2017, of which 3037 were excluded as duplicates, 2877 studies were excluded on title/abstract, and 157 studies were excluded at the full text screening stage. Of the 396 studies included in the EGM by Miller et al. (2020), 40 presented empirical evidence on the impacts of agroforestry interventions, of which only eight used quantitative impact evaluation methods (Figure 2). The additional search identified three further impact evaluations that we included in this SR. This SR focuses on these 11 impact evaluations. All included studies used quasi-experimental methods, with none using an experimental design. Each article reports on one study with multiple components (treatment groups or outcome variables). Table 4 presents descriptive information on the 11 studies and the full data extraction matrix is provided in Supporting Information Appendix 2.

5.1.2 | Included studies

Included studies examined all of the six intervention types described above (Table 2). The most studied interventions were farmer capacity

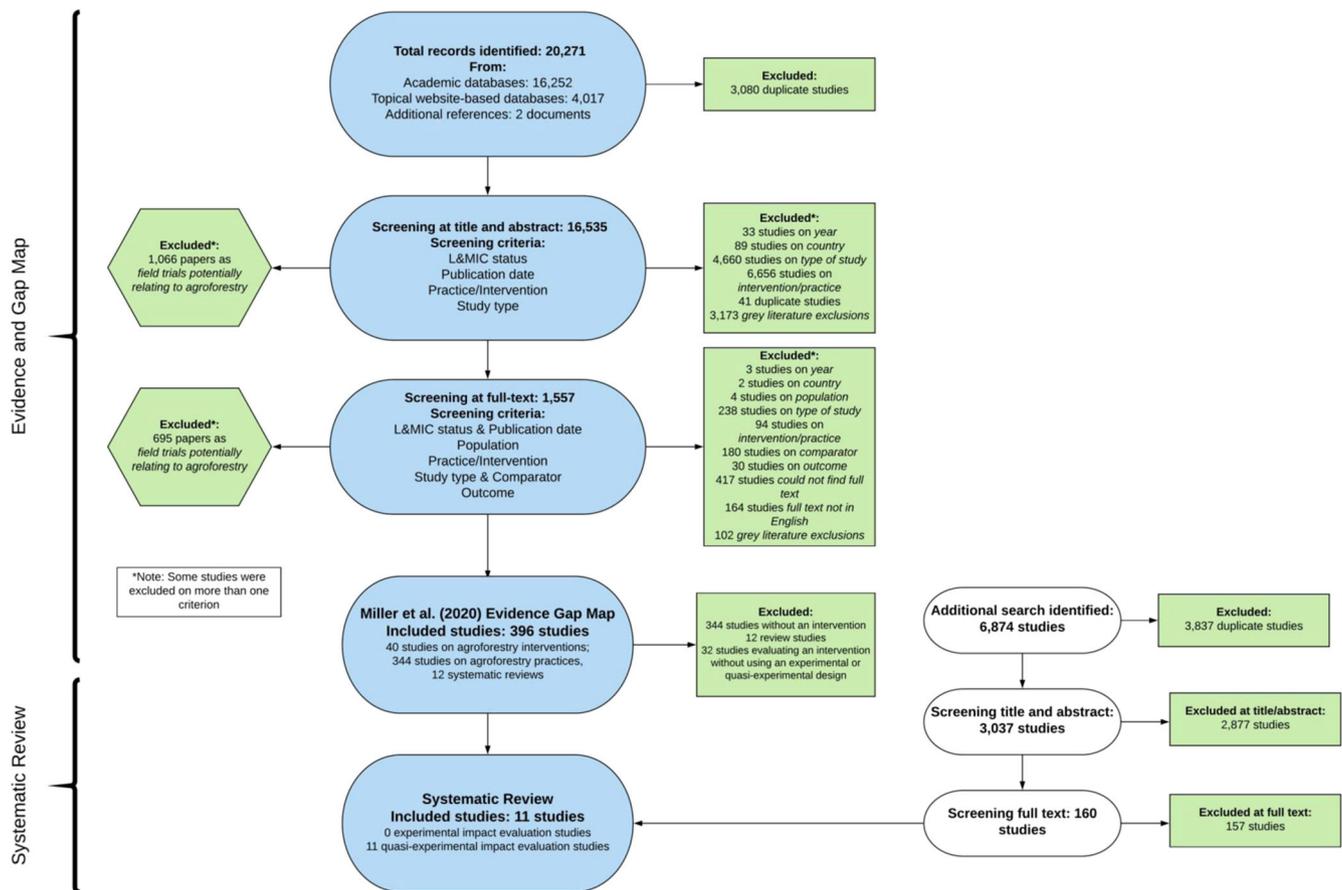


FIGURE 2 PRISMA flow diagram

development ($n = 8$, 73%), of which five (45%) included provision of tree germplasm/seedlings as a component of the intervention, and farmer incentive provision ($n = 3$, 27%).¹ One study (9%) evaluated an institutional and policy change intervention. One study (9%) included a component of community-level campaigning and advocacy, and used social network analysis to understand dissemination of information through different sources to understand adoption as a result of this component of the intervention. One study (9%) included a component of market linkage facilitation.

Nearly all the agroforestry practices promoted in the intervention studies were agrisilvicultural ($n = 7$, 64%) or agrosilvopastoral ($n = 3$, 27%), with one study evaluating a silvopasture intervention (9%). Table 4 shows the specific practices promoted, with trees integrated in crop fields ($n = 5$, 45%) followed by integrated production of animals, crops, and wood ($n = 3$, 27%) the two most frequently promoted.

Ecosystem services was the least frequent outcome category ($n = 4$, 36%), and the most frequent one was human well-being outcomes ($n = 10$, 90%). Agricultural productivity was evaluated for over half of the studies ($n = 6$, 55%). We can see that for specific outcomes, income and household expenditure was the most common outcome

($n = 8$, 73%), followed by agricultural productivity ($n = 6$, 55%). When looking at the combination of interventions and outcomes, the most studied linkages were studies focused on incentive provision and farmer capacity development with human well-being and agricultural productivity outcomes.

These impact evaluation studies were published from 2005 to 2017, with 2016 and 2017 the only years with more than one study appearing. In some years, no impact evaluation studies on this topic were published. All included intervention studies were published in peer-reviewed journals ($n = 9$) except for two (Pender et al., 2008; Place et al., 2005), which were organization reports.

The intervention studies included in this EGM are distributed across several tropical LMICs (Figure 3), with Sub-Saharan Africa having the most countries with a study ($n = 6$, 55%). There were three studies (27%) conducted in Latin America and the Caribbean, and two studies (18%) conducted in East Asia and Pacific. Kenya was the only country where more than one study was conducted ($n = 3$, 27%). The countries with impact evaluations included: Brazil, China, Colombia, Indonesia, Kenya, Malawi, Mozambique, Nicaragua, and Zambia. Almost all fall within the tropics, but there was one study in a temperate climate. Of these study areas, two countries (Malawi and Mozambique) are classified by the World Bank as low-income economies, four countries (Indonesia, Kenya, Nicaragua, and Zambia) are classified as

¹We note that the total percentage here, and at different points throughout this report, can be more than 100% as a given study could include more than one intervention and outcome. Percentage = # of studies meeting criterion/total number of studies ($n = 11$).

TABLE 4 Description of intervention studies using quasi-experimental methods

#	Reference	Title	Pub. year	Country region	Intervention(s) ^a	Practices	Study design	Sample size	Outcomes	Equity measure
1	Bostedt et al. (2016)	Agroforestry extension and dietary diversity—an analysis of the importance of fruit and vegetable consumption in West Pokot, Kenya	2016	Kenya Sub-Saharan Africa	Farmer capacity Enhancing access to tree germplasm	Agrosilvopastoral Integrated production of animals, crops, and wood	Heckman two-stage probit regression	291 farmers (164 received advice, 127 did not)	Food security and nutrition	None
<p>Description: Vi Agroforestry is a Swedish NGO working with tree planting and agroforestry education in Kenya, Rwanda, Uganda and Tanzania. Vi Agroforestry demonstrated and carried out intensive extension work on agroforestry and enclosures. Free advice on agroforestry, including cultivation of trees, crops, fruit and vegetables were given throughout, but provision of tree seedlings and tree seeds (mainly nonfruit trees), changed over time. They use a 2-stage model: selection model (for received advice from Vi Agroforestry) then econometric model for dietary outcomes (effect size on dietary diversity score, number of food groups consumed, and effect size of each specific food group)</p>										
2	Coulbaly et al. (2017)	Adoption of agroforestry and the impact on household food security among farmers in Malawi	2017	Malawi Sub-Saharan Africa	Farmer capacity development Enhancing access to tree germplasm	Agrisilvicultural Trees integrated in crop fields	Endogenous switching regression Instrumental variables	338 farmers (135 adopters, 203 nonadopters)	Productivity (yield) Profitability Income and household expenditure Food security and nutrition	Gender Socio-economic level Literacy/ educational level
<p>Description: This study evaluates an ICRAF agroforestry project in Malawi that promoted fertilizer trees through extension and enhanced access to tree germplasm and nursery establishment as well as farmer training. They used a two-stage model: selection model (measuring the impact of the program on adoption) and then econometric model for impact of adoption on the value of food crops. Study disaggregates by landholding size</p>										
3	Dai et al. (2017)	Assessing the Effect of a Crop-Tree Intercropping Program on Smallholders' Incomes in Rural Xinjiang, China	2017	China East Asia and Pacific	Farmer capacity development Enhancing access to tree germplasm	Agrisilvicultural Trees integrated in crop fields	Propensity score matching	352 households (236 participants, 116 nonparticipants)	Income and household expenditure	None
<p>Description: This study evaluates an agroforestry intervention designed to promote crop-fruit tree intercropping to improve farmer incomes and address scarcity of irrigation water and increasing salinization of arable land in the Xinjiang region, China. They used propensity score matching to estimate the effects of the program on household income</p>										
4	Haggar et al. (2017)	Environmental-Economic Benefits and Trade-Offs on Sustainably Certified Coffee Farms.	2017	Nicaragua Latin America and the Caribbean	Incentive provision	Agrisilvicultural Trees integrated in crop fields	Propensity score matching	278 farms (202 certified farms (see below), 76 noncertified farms)	Reg & Main—Physical, chemical, biological conditions Reg & Main—Mediation of flows Reg & Main—Mediation of waste, toxics and other nuisances Income and household expenditure Productivity (yield)	None
<p>Description: This study evaluates five different Coffee Certification Programs (Organic (n = 47), Utz (n = 35), C.A.F.E. Practices (n = 44), Fairtrade (n = 43), Rainforest Alliance (n = 33)) by matching noncertified farms against the five different types of certification (they measure different impacts for each certification scheme). They measure productivity, costs, and net income. The data were made available to us by the lead author by request</p>										

TABLE 4 (Continued)

#	Reference	Title	Pub. year	Country region	Intervention(s) ^a	Practices	Study design	Sample size	Outcomes	Equity measure
5	Hegde and Bull (2011)	Performance of an agro-forestry based Payments-for-Environmental-Services project in Mozambique: A household level analysis	2011	Mozambique Sub-Saharan Africa	Incentive provision	Agrisilvicultural Trees integrated with plantation crops Trees integrated in crop fields	Propensity score matching	290 households (105 PES, 185 non-PES)	Productivity (yield) Housing and material assets Income and household expenditure	Gender Socio-economic level
<p>Description: This study evaluates a Payments-for-Environmental-Services project in Mozambique, which was an agroforestry-based carbon sequestration project using payments for environmental services. They measured household cash income, consumption expenditure, forest resource use, and crop yields. They used 4 different matching methods, each with a similar but different effect size. The study disaggregates on men versus women headed households and on poor versus rich households</p>										
6	Kuntashula and Mungatana (2013)	Estimating the causal effect of improved fallows on farmer welfare using robust identification strategies in Chongwe, Zambia	2013	Zambia Sub-Saharan Africa	Farmer capacity development	Agrisilvicultural Improved or rotational fallow	Propensity score matching Endogenous switching regression	324 households (113 adopters, 213 nonadopters)	Productivity (yield) Profitability Income and household expenditure Food security and nutrition	None
<p>Description: This study evaluates the impacts of the Kasisi Agricultural Training Centre (KATC), which was an agroforestry extension service promoting improved fallow in Zambia. They measure crop income per household (HH) unit, maize income per HH unit, total maize yield, maize yield per HH unit, maize yield per ha, and months per year with enough grown food. They used 2 different matching methods, each with a similar but different effect size, then used endogenous switching regression for another different effect size set. This study assumes all farmers who adopt did so through the influence of KATC, and only looks at areas where KATC was active (does not clearly measure the impact of the intervention, only the impact of adoption that was attributed to the intervention)</p>										
7	Pagiola et al. (2016)	Evaluation of the Permanence of Land Use Change Induced by Payments for Environmental Services in Quindío, Colombia	2016	Colombia Latin America and the Caribbean	Incentive provision Farmer capacity development	Silvopastoral Trees/shrubs on pasture	Difference-in-differences	98 households (69 PES, 29 non-PES)	Reg & Main—Physical, chemical, biological conditions	Socio-economic level
<p>Description: This study evaluates a Payments for Environmental Services (PES) silvopasture project where the FAO and others provided payments for environmental services for silvopastoral practices in Colombia, plus technical assistance in some cases. They measure “ESI Generation” for 4-year PES, 2-year PES with technical assistance, and 4-year PES with technical assistance. “ESI Generation” is the before/after change in “environmental services index” (ESI), which aggregates indices of the biodiversity conservation and carbon sequestration services associated with specific silvopastoral practices. They measured the impacts both during project and post-PES, with different effect size for each period</p>										
8	Pender et al. (2008)	Impacts of the Hutan Kamsyarakatan social forestry program in the Sumberjaya watershed, West Lampung district of Sumatra, Indonesia	2008	Indonesia East Asia and Pacific	Institutional and policy change	Agrisilvicultural Trees integrated in crop fields	Propensity score matching Instrumental Variables	640 plots (160 HKm, 101 HKm pending, 160 without HKm, 160 park/private)	Housing and material assets Income and household expenditure	None

TABLE 4 (Continued)

#	Reference	Title	Pub. year	Country region	Intervention(s) ^a	Practices	Study design	Sample size	Outcomes	Equity measure
<p>Description: This study evaluated the Hutan Kamasyarakatan social forestry (HKM) permitting program in Indonesia. The HKM program provides groups of farmers with secure-tenure permits to continue farming on state Protection Forest land in Sumatra in exchange for protecting remaining natural forestland, planting multi-strata agroforests, and using recommended soil and water conservation (SWC) measures on their coffee plantations. They measured impact on land value, investment in soil and water conservation practices and soil fertility management, and profitability</p>										
9	Place et al. (2005)	The impact of agroforestry-based soil fertility replenishment practices on the poor in western Kenya	2005	Kenya Sub-Saharan Africa	Farmer capacity development Community-level campaigning and advocacy	Agrisilvicultural Improved or rotational fallow	Instrumental Variables	103 households within pilot villages with adopters and nonadopters	Productivity (yield) Food security and nutrition Income and household expenditure Housing and material assets	Gender economic level Race/ethnicity Literacy/educational level
<p>Description: This study evaluates the impacts of an IFPRI program promoting soil fertility replenishment (SFR) practices, including improved fallow. IFPRI provided extension to promote and assist in the establishment of improved fallows in Kenya. They used a two-stage model: selection model (measures impact of program on adoption) and econometric model measuring the impact of adoption on outcomes. They measured liquid asset change, changes in nonfood expenditures, changes in per capita nonfood expenditures, and changes in food intake and nutritional status for econometric analysis of the adoption of improved fallows. They also show results of farmer-managed field trials. They measured seasonal per acre net gain in profitability, return to labor, and yield for field trials</p>										
10	Sills and Harris (2015)	Evaluating the long-term impacts of promoting "green" agriculture in the Amazon	2015	Brazil Latin America and the Caribbean	Farmer capacity development Enhancing access to tree germplasm Market linkage facilitation	Agrosilvopastoral Integrated production of animals, crops, and wood	Propensity score matching Difference-in-difference	Varied by year (total sample ranges 195-528 households, 36-64 matched)	Maintenance of physical, chemical, biological conditions Income and household expenditure Housing and material assets	None
<p>Description: This study evaluates the impacts of participating in the Alternative Producers Association (APA) in Brazil on the short- and long-term outcomes of the program, including diversification and agroforestry, deforestation level, and household income and wealth. They use propensity score matching and difference-in-differences to evaluate the effects of the program</p>										
11	Thorlakson and Neufeldt (2012)	Reducing subsistence farmers' vulnerability to climate change: evaluating the potential contributions of agroforestry in western Kenya	2012	Kenya Sub-Saharan Africa	Farmer capacity development Enhancing access to tree germplasm	Agrosilvopastoral Integrated production of animals, crops, and wood	Propensity score matching	116 households (73 participants, 43 nonparticipants)	Productivity Profitability Income and household expenditure Mediation of flows Energy provisioning	Gender
<p>Description: This study evaluates the impacts of an ICRAF program promoting agroforestry techniques to help subsistence farmers reduce their vulnerability to climate change in Kenya. They use propensity score matching to compare farmers engaged in the agroforestry project with a control group of neighboring farmers to estimate the effects on farm productivity, off-farm incomes, wealth and the environmental conditions of their farms</p>										

^aWhen more than one intervention type is listed, this indicates that there were multiple components to the intervention recorded.

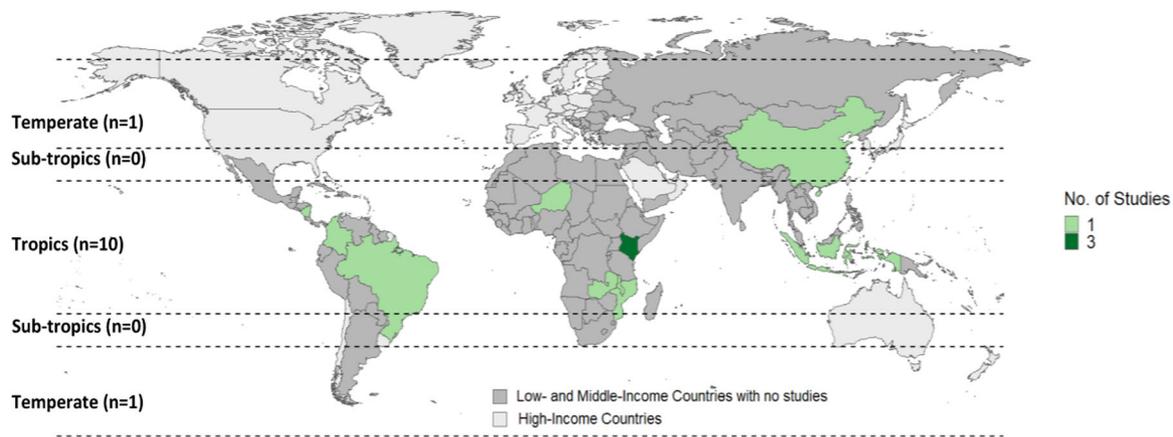


FIGURE 3 Distribution of intervention impact evaluation studies by country climatic zone

lower-middle income economies, and three countries (Brazil, China, Colombia) are classified as an upper-middle income economy. In Figure 3, countries in dark grey are LMICs where we found no relevant studies on agroforestry interventions.

Five of the eleven included impact evaluations (45%) presented results disaggregated by at least one measure of equity (Table 4). Three of these five studies presented results disaggregated by two or more measures of equity, one of which presented results disaggregated by all four measures captured in this EGM. The most common groups disaggregated in the agroforestry intervention studies were socioeconomic level and educational level ($n = 4$ each) with three studies disaggregating results by gender and one by race/ethnicity.

5.1.3 | Excluded studies

From the agroforestry EGM, 388 studies were excluded for this SR because the study design did not meet the requirements for inclusion in our SR, namely they did not consider an intervention, as opposed to a practice only, and did not use experimental or quasi-experimental evaluation methods. From the additional search conducted after the EGM, 157 studies were excluded at the full-text level for the same reasons, that is, no intervention or not using experimental or quasi-experimental methods.

5.2 | Risk of bias in included studies

The results of the risk of bias assessment are summarized in Figure 4. The overall risk of bias assessment results are shown in Figure 5. Two of the studies were also included in Snilstveit et al. (2019), and we used their assessments of risk of bias for those two studies. The full risk of bias assessment can be found in Supporting Information Appendix 3.

Most of the studies only discussed selection bias or did not discuss the risks of bias at all. Only two studies attempted to rigorously address selection bias through the use of endogenous switching regression (Coulibaly et al., 2017; Kuntashula & Mungatana, 2013), but there are

still risks of selection bias using these techniques and there is no guarantee of group equivalence. In these two studies, the authors explained how quasi-experimental methods allowed them to account for selection bias by using endogenous switching regression to account for endogeneity bias and the effects of unobservable covariates. One other study also used an endogenous selection model, but only education level and two of the villages were used as variables in the selection equation (Bostedt et al., 2016). The other studies use PSM or IVs, resulting in varying degrees of potential bias. All of the studies had self-selection into the treatment group, which makes it very unlikely that the selection bias was completely controlled through matching techniques. Most of the studies using matching techniques based on only on end line data rather than baseline data, and almost all were cross-sectional studies of end line data. Only one study evaluated households at multiple points in time, including baseline and end line data (Pagiola et al., 2016).

While few studies explicitly discussed any form of bias, we found little risk of performance bias and outcome measurement bias. Analysis in all included studies was based on farm or household survey data that were consistently applied across treatment and control groups and there was little concern about risk of performance bias between the groups. Only two studies were coded as unclear for risk of performance bias. We found two studies that were unclear if they were at risk for outcome measurement bias; otherwise, we found no evidence of outcome measurement bias since there were no incentives for the groups or the enumerators to exaggerate their responses and the surveys were consistently implemented and timed.

Most of the studies did not explicitly discuss potential spillover effects. All of the studies selected treatment and control groups from the same population, not geographically separated. Studies with low risk of spillover bias clearly defined the cutoff between treatment and control groups, for example, minimum number of years farmers had to have adopted improved fallow for effects to be observable as cutoff for treatment and control groups. Other studies had a high risk of spillover bias since they examined the impacts of directly receiving extension advice, which could easily be transferred to friends and neighbors. For certifications schemes and PES programs, there could be some

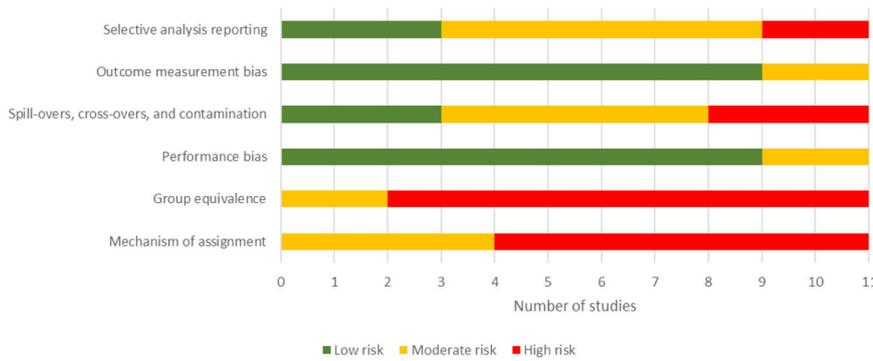


FIGURE 4 Summary of risk of bias assessment results of eleven impact evaluations

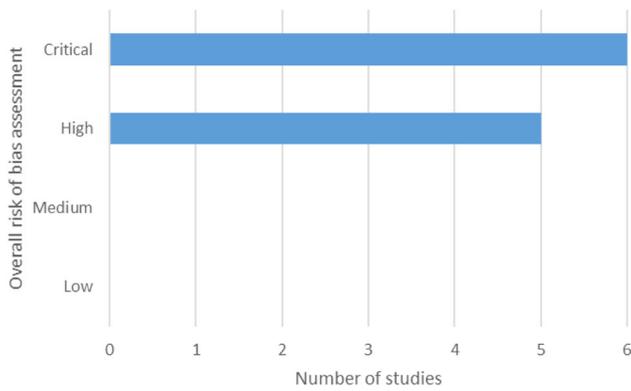


FIGURE 5 Summary of overall risk of bias assessment

contamination of the control groups if they were anticipating adopting the program. Although most of the studies did not describe the potential for spillover bias, it is likely to exist since the practices are visible and incentives available to adopt the practices. Since the studies were conducted within the same population (adopters or program participants vs. nonadopters or nonparticipants within the same area), spillover bias would bias the effect sizes.

There were some concerns over selective analysis reporting, particularly about reporting tests verifying the methods of analysis. Many of the studies reported results based on multiple analysis methods and reported all statistically significant and insignificant

results for all outcome measures discussed, which implied a lower risk of bias due to selective analysis reporting.

The overall body of evidence has a high risk of bias due primarily to self-selection into studied interventions. All of the studies had a high risk of bias due to issues with self-selection into the program and analysis based only on cross-sectional data. We highlight that the evidence base did not include any RCTs, an important method for addressing this potential source of bias.

5.2.1 | Synthesis of results

Effects of agroforestry interventions on yields

We identified six studies that measured the effects of agroforestry on agricultural productivity in terms of yields. We performed a meta-analysis across five of these studies for the effect size of agroforestry interventions impacts in terms of percentage increase or decrease in productivity for the treatment group against the control group. We report the summary effect size as described in Section 4.3.5 for the Haggar et al. (2017) study, which reports results for five different treatment group compared to the same control group, as well as for the Thorlakson and Neufeldt (2012), which reports results for two subregions. Results are presented as a forest plot in Figure 6.

The average effect of these interventions on crop yield outcomes is 1.16 standardized mean difference (Hedges' *g*) with a 95% confidence interval (CI) of [-0.35, 2.67] ($p = .13$), which we calculated using a random effects model. The forest plot in Figure 6 shows a

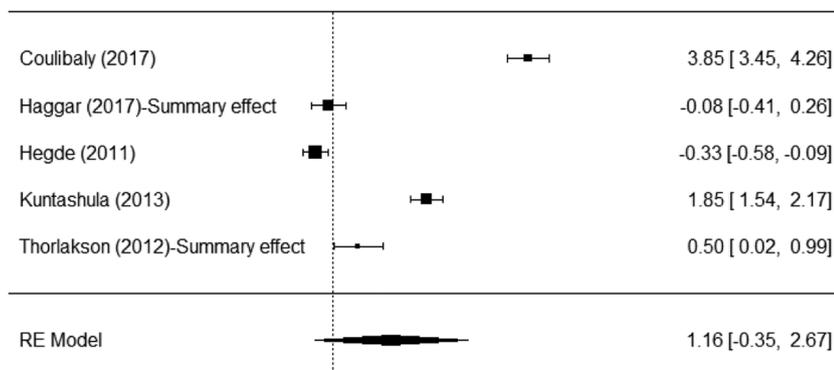


FIGURE 6 Forest plot of yield effect size (standardized mean difference, Hedges' *g*), where positive indicates yields increased, negative indicated yields decreased

high amount of variability among studies, which is supported by the statistical heterogeneity tests ($I^2 = 98.99\%$, $\tau^2 = 2.94$, $Q(df = 4) = 370.7$). Effect sizes range from -0.33 SMD (95% CI $[-0.58, -0.09]$) for a PES program in Nicaragua (Hegde & Bull, 2011) to 3.85 SMD (95% CI $[3.45, 4.26]$) for the effect of the improved fallows on crop value (a proxy for crop yield) in Malawi (Coulibaly et al., 2017).

The overall effect size is not statistically significant, but there was high variability in the contexts and types of practices. Several studies found large, positive effect sizes, all of which evaluated the impacts of soil fertility replenishment practices and climate change resilience practices in Sub-Saharan Africa (Coulibaly et al., 2017; Kuntashula & Mungatana, 2013; Thorlakson & Neufeldt, 2012). On the other hand, the studies that employed incentives to promote adoption of agroforestry practices for biodiversity conservation and carbon sequestration resulted in reduced yields (Haggar et al., 2017; Hegde & Bull, 2011). Given the low number of included studies, we cannot identify specific contextual factors that would generalize as key factors driving yield effect direction or size due to agroforestry adoption within this study.

Included studies varied widely in the methods used, agroforestry intervention and system studied, and outcomes reported. Table 4 describes some of these important differences, which we further discuss below.

One of the studies, Haggar et al. (2017), evaluated five different coffee certification schemes in Nicaragua designed to promote environmentally-friendly, shade-grown coffee production practices. Each of these schemes was evaluated against a common control group, so we used the summary effect size as described in Section 4.3.5. The study matched farms certified under each program with noncertified farms. The study design implied that noncertified farms did not use agroforestry practices and thereby serve as a relevant control; however, based on available information in the study, it was not clear that there were in fact no agroforestry practices used on noncertified farms. Two of the certification schemes, Organic and Utz, resulted in statistically significant reductions in productivity (Hedge's $g = -0.488$ (0.043) and -0.465 (0.057), respectively), while the other three, C.A.F.E. Practices, Fairtrade, and Rainforest Alliance resulted in insignificant increases in productivity (Hedge's $g = 0.337$ (0.045), 0.151 (0.046), and 0.088 (0.059), respectively). The authors conclude that the outcomes of different coffee certification schemes were highly variable based on the requirements established by the certification program. For example, they noted that there were pre-existing differences among farmers under difference certification schemes due to differences in program requirements. There were also differences in types of farmers who were targeted for the programs. NGOs and social enterprises promoted Fairtrade and organic certifications and targeted smaller-scale, more disadvantaged farmers. The other three schemes, C.A.F.E. Practices, Rainforest Alliance, and Utz-certified, were promoted through coffee traders and targeted medium- to large-scale enterprises. They also found that tree diversity and carbon stock tradeoffs with productivity were mediated by level of investment in labor and inputs, where farmers with higher tree

diversity invested less, had lower productivity, but received a premium price. The certification schemes that required lower investment in labor and inputs tended to have lower productivity and those with higher investment in labor and inputs tended to have higher productivity. However, the net revenue did not necessarily differ due to the lower investment costs and higher price premiums. The requirements on environmental and social criteria vary substantially between the different certification schemes, which also results in variability in the distribution and diversity of shade trees on the coffee farms. Overall, the authors concluded that the certification schemes were delivering enough compensation to offset the lower returns on investment.

Hegde and Bull (2011) evaluated the crop yield impacts of a PES program supporting the planting of trees on farms (along field boundaries or intercropped) for carbon sequestration in Mozambique. A variety of crops were planted by the mainly subsistence farmers in the study area, so the authors used crop value as a proxy measure for crop productivity. They report a large decrease in crop value, indicating a significant loss in crop yields. The effects on agricultural crop value were of a similar magnitude and direction for poor households, but the poor households felt a significantly more negative effect on the value of forest products. The crop value effects were more acute for women-headed households, which had larger decreases in crop value under the PES scheme than male-headed households.

Coulibaly et al. (2017) measured the impact of fertilizer tree adoption on food crop value, defined as the yield of all food crops per year multiplied by the farm gate price for each food crop that year. They used an IV approach along with endogenous switching regression to show the impact of the intervention on adoption and the impact of adoption on food security outcomes. The instruments they used were participation in agroforestry training, which they argued affected the decision to adopt but did not affect the dependent variable except through adoption. They tested this instrument and found that it was positively correlated with the decision to adopt agroforestry but was not significantly related to food productivity. However, it is likely that there are systematic differences between farmers who decide to participate in agroforestry training programs and those who do not. They find a significant, large, and positive impact on crop value (35% increase). They also examined the heterogeneity in their results, and they found that households with land holdings less than two acres experience the greatest benefits from adopting fertilizer trees. Farmers with land ownership of less than one acre averaged an 82% increase in the food crops value from adopting fertilizer trees, and farmers with between one and two acres averaged a 66% increase in food crop value with fertilizer tree adoption. They also measured maize yields, but they did not use quasi-experimental methods to analyze the results.

Kuntashula and Mungatana (2013) evaluated the impact of improved fallows in Zambia on total maize yield of the farm, on maize yield per person in the household, and on maize yield per hectare. For our meta-analysis, we look at total maize yield (ton) per hectare since this was the most similar outcome measure to those of the other included studies. Using an endogenous switching regression model,

they found highly heterogeneous effects on outcome measures. The average treatment effects on the untreated (ATU) are significantly higher than the average treatment effect on the treated (ATT) outcomes, which implies that the farmers who did not adopt are the ones that would have received the higher benefits from adopting the improved fallow technology. Importantly, they found that increases in maize yields were lower in their study with a quasi-experimental design compared to increases in maize yields observed in randomized field experiments conducted by ICRAF elsewhere in Zambia (Mafongoya et al., 2006). The authors attribute these differences to farmers' skills in managing improved fallows and maize crops, highlighting the need for continuous farmer training in new agricultural technologies such as improved fallows, and demonstrating potential differences between controlled field experiments and larger-scale intervention outcomes.

Thorlakson and Neufeldt (2012) evaluated the impacts of an agroforestry intervention that provided farmer training, seedling provision, and small amounts of food for participating as well as training, tools, and seedlings for tree nursery management. They measured farm productivity by converting current seasonal crop production to economic units using average 2010 crop prices in the study region. Through quantitative and qualitative methods, they found that farmers tended to have increased yields with agroforestry practices from fruit production, though they noted that the full benefits may not have been realized within the 4 years since the start of the program. The increase in productivity in the Lower Nyando region was approximately a 35% increase, and in the Middle Nyando region was approximately a 20% increase, but the effect sizes in both regions was not statistically significant.

Finally, we note that Place et al. (2005) included results of improved fallow on yields in Kenya. Due to insufficient information reported and since this was not part of their econometric analysis, this study was not included in the meta-analysis. In their qualitative analysis, the farmers reported perceptible increases in yields that they directly attributed to the improved fallow technology. Based on farmer recall data, the study reports a median increase in maize yield of 167% and a mean increase of 128% compared to an unfertilized maize-only control, but they did not report standard errors or p-values and they reported that 12.5% of the plots had a decrease in maize yield.

Results of the foregoing meta-analysis show there is high variability in both the magnitude and direction of effect sizes of the impact of agroforestry interventions on productivity. This result is not surprising. Agroforestry encompasses a wide array of practices, some of which are directed at conservation as the primary objective, such as coffee agroforests and trees for carbon sequestration, and others are directed at soil fertility replenishment, such as improved fallows and fertilizer trees. Here, we observe that there is not only an apparent divide in the types of agroforestry practices and their objectives but also in the types of interventions used to support the adoption of those practices.

In the two studies (Hagggar et al., 2017; Hegde & Bull, 2011) where we observe declines in yields, interventions were designed to counteract negative yield effects with higher prices or PES. Among

the different coffee certification programs, some resulted in higher yields than their matched noncertified farms along with receiving higher prices for their products through the certification program. The intention of the coffee certification programs was to promote the maintenance of sustainable practices, such as shade-grown coffee agroforestry systems, to provide habitat for biodiversity conservation. The interventions were designed to incentivize farmers to maintain these practices by increasing crop market value due to the certification to counteract any losses in yields or increased difficulty in managing the system. Results were inconclusive on the yield side, as described above. The PES program in Mozambique promoted the single-purpose planting of trees for carbon sequestration, instead of multi-purpose trees that could be productive in their own right. A distinct, drastic loss of yields was indeed observed, with the program intended to deliver direct payments to farmers to counteract these losses (see results on income below).

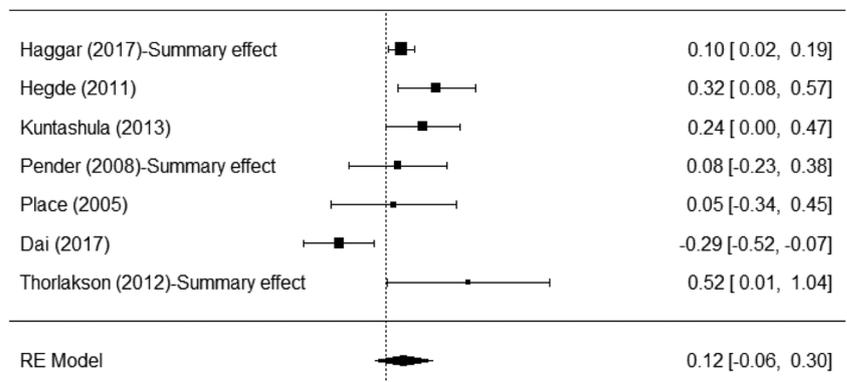
On the other hand, in the four cases where we observe increases in yields (Coulibaly et al., 2017; Kuntashula & Mungatana, 2013; Place et al., 2005; Thorlakson & Neufeldt, 2012), the interventions were meant to increase farmer capacity through the provision of extension and training services to help farmers adopt practices that would ultimately improve productivity. These practices included integrated production systems, improved fallows, and fertilizer trees. We note that all four of these programs were implemented in Africa, in locations with severe soil fertility limitations. Such low baseline conditions combined with interventions tailored to boost agricultural productivity appear to have been propitious for increasing yields.

Effects of agroforestry interventions on income

We identified eight studies that measured the effects of agroforestry on income. We conducted a meta-analysis based on seven of these included studies that measured the impacts of agroforestry on income. Again, we report the summary effect size as described in Section 4.3.5 for the Hagggar et al. (2017) study and the Pender et al. (2008) study, which report results for multiple different treatment groups compared to the same control group and for Thorlakson and Neufeldt (2012), which reports multiple comparisons. Results are presented as a forest plot in Figure 7.

The average effect of these interventions on household income outcomes is 0.12 standardized mean difference (Hedges' g) with a 95% CI of $[-0.06, 0.30]$ ($p = .20$), which we calculated under a random effects model. This result indicates that there is an overall small, positive, but statistically insignificant, effect of agroforestry interventions on incomes. In aggregate, then, studied agroforestry interventions have improved incomes either through improved yields or through PES or higher prices received by eco-certification. The forest plot in Figure 6 shows moderately high variability between studies, which is supported by the statistical heterogeneity tests ($I^2 = 75.29\%$, $\tau^2 = 0.04$, $Q(df = 6) = 19.16$). The effects range from -0.29 SMD (95% CI $[-0.52, -0.07]$) for the effect of crop-fruit tree intercropping in China (Dai et al., 2017) to 0.52 SMD (95% CI $[0.01, 1.04]$) for the effect of the agroforestry program to reduce vulnerability to climate change in Kenya (Thorlakson & Neufeldt, 2012).

FIGURE 7 Forest plot of income effect size (standardized mean difference, Hedges' g), where positive indicates incomes increased, negative indicated incomes decreased



As with the studies evaluating the impacts of agroforestry on yields, there was high variability in the outcome measures evaluated in each study. Some of the specific differences between studies as well as heterogeneity and equity effects are discussed below.

Kuntashula and Mungatana (2013) evaluated the impact of improved fallows in Zambia on crop income, generally, and maize income, specifically, per person (in man equivalent units) at the household level. Our meta-analysis examined crop income per person in the household since this was the most similar outcome measure to those of the other included studies. Although maize income significantly increased due to the improved fallow practices, this did not translate to significant increases in total crop income. Part of the reason for this result is that farmers used improved fallow primarily for maize fields and not for high-value crop fields, limiting the ability of improved fallows to contribute to total crop income. Results exhibited high levels of heterogeneity, however. The ATU effects were significantly higher than ATT outcomes, which implies that the farmers who did not adopt were the ones that would have received the higher benefits from adopting the improved fallow technology.

Place et al. (2005) performed an econometric analysis using IVs for the impacts of improved fallow on liquid assets change and change in nonfood expenditures in Kenya. Our meta-analysis includes the effects of improved fallow adoption on liquid asset change, but the authors also evaluated the effect on changes in nonfood expenditures using econometric techniques. We use liquid asset change in the meta-analysis as this is the closest measure of net income; however, while the results of liquid asset change were positive but nonsignificant, the changes in nonfood expenditures were negative and significant. They also reported on farmer-managed field trials. In the farmer-managed field trials, a positive seasonal per acre net gain and a positive return to labor for *tephrosia* fallow and *crotalaria* fallow were observed compared to the no-input control case.

Thorlakson and Neufeldt (2012) measured the impact of the agroforestry program on household wealth, using an estimate of total livestock value as a proxy since livestock is the most frequently cited indicator of farmer wealth among the tribes studied here in western Kenya. There were differences in effects between the subregions, with Lower Nyando having a positive, statistically significant impact on wealth and Middle Nyando having a negative, statistically insignificant effect on wealth. Wealth effects were strongly associated with length

of time since the farmers planted the trees. Approximately 87.5% of farmers in the Lower Nyando who had their trees for 4 years reported income improvements, but most of the farmers in Middle Nyando had their trees for only 2 years, which was not enough time for them to become productive and generate additional income.

The study by Haggar et al. (2017) of the five different coffee certification schemes in Nicaragua measured impacts on cost of production in US\$ per hectare and net income in US\$ per hectare. The found variability again in these results based on the scheme evaluated. C.A.F.E. Practices and Fairtrade both were associated with increases in costs and increases in net income (for net income, Hedge's g = 0.505 (0.046) and 0.375 (0.046), respectively). Organic and Utz certification, on the other hand, both showed decreases in cost and decreases in net income compared to their matched noncertified farms (for net income, Hedge's g = -0.031 (0.042) and -0.558 (0.058), respectively). Finally, for Rainforest Alliance, there was a decrease in costs and an increase in net income, but neither of these results showed statistical significance at 10% level (for net income, Hedge's g = 0.351 (0.060)).

Hegde and Bull (2011) evaluated the impacts of on a PES program supporting the planting of trees on farms (along field boundaries or intercropped) for carbon sequestration in Mozambique on cash income per capita and expenditure per capita. They find a significant increase in both expenditure per capita and cash income per capita, indicating a significant gain in income and expenditure. For women-headed households and poor households, however, these gains were not observed. Women-headed households had a decrease in expenditure per capita and much lower magnitude increase in cash income per capita than the full population. Poor households had a small increase in expenditure per capita and in cash income per capita. These results for women-headed and poor households were much smaller than for the general population and not statistically significant from zero.

Pender et al. (2008) measured the impact of a social forestry program in Indonesia (Hutan Kamasyarakatan—HKm), which provided secure tenure permits in exchange for incorporating soil and water conservation measures. They considered two treatment groups: those with an HKm permit and those with an HKm permit pending, which were compared to the same control group. The effect sizes for these treatment groups were combined into a summary effect size, per Section 4.3.5. The outcome measures related to land

value and profitability. For land value, they used an IV approach, finding a large, positive, but not statistically significant effect size. For profitability, results were mixed, with a negative, insignificant effect between permit and pending permit groups and between permit and no permit groups, but a positive, significant effect between pending permit and no permit groups (Hedge's $g = -0.024$ (0.008) for permit compared to no permit and Hedge's $g = 0.181$ (0.008) for permit pending compared to no permit). HKm had mixed impacts on profits, with timber trees reducing profitability because timber harvesting was not allowed and multipurpose nontimber trees contributing to increased profits. The estimated additional profit resulting from additional multipurpose trees as a result of an HKm permit represented a significant increase in farmers' incomes resulting from an HKm permit. They also found a perceived increase in tenure security because of the HKm program of 26.4% (SE: 2.0).

Dai et al. (2017) assessed the impact of a crop-fruit tree intercropping program on participants' gross, on-farm, and off-farm incomes in rural Xinjiang, China. They found a negative effect of the program on incomes since the additional time needed to manage the agroforestry systems reduced off-farm income opportunities. We included the gross income effects in the meta-analysis to be consistent with the other included studies. Land tenure security was determined to be major factor in the negative effects of this program since farmers in the region were reluctant to invest the time and resources in the management the agroforestry systems when they did not have tenure security. There were also issues with crop yield losses due to tree-crop competition for irrigation water, light, and fertilizer. The authors suggested that the program was poorly targeted and implemented, with difficulties and limitations in obtaining the subsidies, insufficient subsidy amounts, and low tree survival rates.

Another study, Sills and Caviglia-Harris (2015), also evaluated the impacts of a program promoting "green" agriculture in the Brazilian Amazon in the short- and long-terms. We did not include the results from this study in this meta-analysis since they did not observe the impacts of the agroforestry component directly, but rather the impacts of the program on agroforestry adoption and the overall impacts of the program, which included a broad range of green agricultural practices. The program increased agroforestry adoption in participants. In the short-term, they found positive, significant impacts of the program on household income, but mixed insignificant impacts in the medium- and longer-terms. We tested the impact of including these results, and there was no substantive change in our meta-analysis results with included this result.

We found that in the first three cases in Africa, where improved fallow, intercropping, or fertilizer trees were the promoted practice, there was evidence that adopting agroforestry improved or diversified yields and led to increased incomes for farmers. While the effects were positive, the results were not consistently statistically significant. These interventions fall under the category of providing support and technical assistance to farmers to adopt profitable agroforestry practices. The last study discussed, Dai et al. (2017), was expected to generate improved and diversified income, but demonstrates that contextual factors such as land tenure, soil and water limitations, and

opportunity costs as well as the program implementation strategy can hinder the effectiveness of an agroforestry program.

In the other three cases, we saw examples of PES, certification schemes, and land tenure security programs, which all fall under the category for incentive provision interventions. In these cases, there was variability in the productivity and profitability of the practice, but the provided incentives were meant to lead to increased incomes. For the coffee certification schemes, there was high variability between the schemes. Incomes were particularly affected by whether the farmers received higher prices for the certified products or not. In the PES case, there was a drastic reduction in yields, as we saw in the previous section, but the farmers had a significant increase in income and expenditure per capita. However, this did not carry over to women-headed households or poor households, which saw lower increases in income and expenditure. The PES payouts were mostly distributed in the first 6 years, with only 10% of the total payment to be delivered at year 25, the term of the agreement. The study was conducted only 4 years after the initiation of the project, so the long-term impacts on income are questionable. In some areas, farmers planted multi-purpose trees that would reach productive maturity around the end of the bulk of the payments. In cases where trees are not productive and they significantly reduce yields, the permanence of these systems and project attrition rates after year 6 are not known.

Effects of agroforestry interventions on nutrition and food security

We found four studies that measured some dimension of nutrition and food security outcomes. The measures they used were not consistent, so we did not perform a meta-analysis for this outcome.

Bostedt et al. (2016) focused on the impacts of agroforestry on nutrition in their study of a Vi Agroforestry project in Kenya. The project provided free advice and technical training on agroforestry, and for a portion of the project provided free tree seedlings. They measured the impact of integrated agrosilvopastoral systems on dietary diversity score and food group consumption. They find a significant, positive effect on their outcome measures with the adoption of agroforestry. They used the Heckman two-stage estimation procedure, with the first stage as the selection model and the second stage estimating the impact of receiving agroforestry advice through the intervention on their outcome measures to compare the treatment and control groups. They find that receiving Vi Agroforestry advice increased the dietary diversity score by 1.22 (t value: 2.19, $p < .05$) and increased the number of food groups consumed by 2.291 (t value: 3.87, $p < .001$), out of the 12 food groups analyzed. For the specific food groups, significant effects were found for food groups B (dark yellow/orange-fleshed roots, tubers and others), C (roots, tubers and plantains), D (dark green leafy vegetables) and F (other fruit/vegetables), with the strongest marginal effect on food group B. The household dietary diversity score is defined as the number of unique foods consumed by household members over a given period, on a scale from 0 to 12. The baseline dietary diversity score for this region was 3.98.

On the other hand, Place et al. (2005) evaluated changes in food intake and nutritional status due to improved fallow agroforestry, but they find no significant results on nutritional measurements. Instead,

the only variable that they found that significantly increased energy, protein, and iron intake was if the household was female headed. In their analysis, female-headed households had significantly positive (or less negative) change in each of these three indicators.

In addition to these two studies that consider nutrition as outcomes, two other studies evaluate the impact of agroforestry on food security. In Coulibaly et al. (2017), food security is measured in terms of food crop value, and they find a strong, positive increase in their study of fertilizer trees in Malawi. They concluded that adopters of fertilizer trees were more food secure due to significantly higher maize yields and food crop values than nonadopters. For the full sample, they find that fertilizer trees increased food crop value by 35%, or 12,447 Malawian Kwacha (MWK) per acre (*SD*: 4615.59, $p < .01$), as discussed previously in the yields section. Disaggregating by landholding size, they find that smallholder farmers receive the highest benefits from adopting agroforestry. Farmers with less than one acre of land received an increase of 82% in food crop value, or 32,433.89 MWK per acre (*SD*: 9795.91, $p < .01$) and farmers with between one and two acres of land received an increase of 66% in food crop value, or 24,261.90 MWK per acre (*SD*: 8896.76, $p < .01$) due to adoption.

Kuntashula and Mungatana (2013) evaluated the impact of improved fallows in Zambia on months per year with enough grown food. Using Kernel matching, they find an average treatment effect on the treated of a 1.069 (*SE*: 0.323, *t* value: 3.311) month increase in number of months per year with enough grown food. Results were similar in their nearest neighbor matching analysis. They conclude that improved fallows increased food security by increasing production of maize and through increased incomes from maize crops.

Across the four studies that looked at nutrition and food security as an outcome measure, there was a positive or neutral impact of agroforestry on nutrition and food security. Although the evidence base is thin, it currently favors the hypothesis that agroforestry can improve nutritional and food security outcomes. That there were so few rigorous studies on this topic is surprising given that agroforestry is often promoted for diversifying food production and thereby improving nutrition and food security outcomes.

Effects of agroforestry interventions on the environment

Five studies measured the effects of agroforestry interventions on environmental indicators. Because these studies used a variety of different indicators, we were not able to perform a meta-analysis.

Pagiola et al. (2016) focused on environmental impacts as the primary outcome of a PES project promoting silvopasture in Colombia. They assessed the effects of silvopasture on an environmental service index (ESI, 0–2 point scale), which aggregates indices of the biodiversity conservation and carbon sequestration services. Overall, the changes due to the project resulted in the ESI/ha of PES recipients increasing by over 60% ($p < .01$). Post completion of PES program, the overall ESI/ha of former PES recipients declined slightly (by <2%), and that of control households increased (by almost 9%) but neither change was statistically significant. Three PES program scenarios were considered: a 4-year PES program, a 2-year PES program with technical assistance, and a 4-year PES program with technical assistance. During the project, the

impacts on ESI/ha was 0.308 (*SE*: 0.186, $p < .10$), 0.343 (*SE*: 0.121, $p < .05$), and 0.549 (*SE*: 0.116, $p < .05$), respectively for each of the three programs. The post-PES changes in ESI/ha were small and insignificant, with values of -0.041 (*SE*: 0.091, NS), 0.052 (*SE*: 0.062, NS), and 0.026 (*SE*: 0.063, NS), respectively for each of the three programs. Therefore, there were persistent, longer-term, positive impacts of the PES program on biodiversity conservation and carbon sequestration outcomes.

Pender et al. (2008) measured the impact of the aforementioned HKm program in Indonesia on tree planting and investment in soil and water conservation practices. They found that the program significantly increased the planting of timber trees by 326.5 trees per hectare (*SE*: 166.7, $p < .05$) and of multipurpose trees by 472.8 trees per hectare (*SE*: 234.6, $p < .05$) with the HKm permit against land without HKm. There was a small, insignificant decrease in shade trees of -41.3 trees/ha (*SE*: 261.1) and not statistically significant increase in planting of coffee of 178.8 trees/ha (*SE*: 1924.0). They also found that land with the HKm application in process had increases in timber and multipurpose tree plantings, though they were lower than those on lands already with the HKm permit. However, they found small, insignificant effects on investment in soil and water conservation practices. Lands with an HKm permit slightly increased their investment in sediment pits and fertilizer and slightly decreased their investment in compost. The researchers conclude that with the planting of timber and multipurpose trees, the program was achieving its soil and water conservation objectives, based on other research by ICRAF and others demonstrating the positive environmental impacts of these types of plantings.

The study by Haggard et al. (2017) in Nicaragua measured impacts on habitat quality, tree carbon-stocks, and soil and water conservation. They found variable results across the different schemes, which each scheme contributing some positive benefit to at least one of these measures, but none of the schemes contributing positively to all of them. Fairtrade, Organic + Fairtrade, Rainforest Alliance, and Utz all had positive impacts on Margalef tree diversity index, which was a proxy measure for habitat quality, while C.A.F.E. Practices was the only scheme with a negative impact on this indicator variable. C.A.F.E. Practices, Fairtrade, Organic + Fairtrade, Rainforest Alliance, and Utz all had positive impacts on above ground carbon stocks, though these results were not statistically significant. They found that the certification schemes positively impacted the use of soil and water conservation methods, such as ground cover, recycling of coffee pulp, and application of organic fertilizers. Organic, Rainforest Alliance, C.A.F.E. Practices and Utz had at least 20% more farms who reduced the volume of water used for coffee processing and had good management of wastewater contaminated from coffee processing or domestic sources compared to noncertified farms.

Sills and Caviglia-Harris (2015) measured the difference in percent of the farmer's lot that was deforested between the "green" agriculture program participants and nonparticipants. They found that the program decreased forest cover loss, but the results were not statistically significant and may have been due to positive selection bias.

Thorlakson and Neufeldt (2012) measured the effects of the agroforestry program in Kenya on soil erosion and tree biomass. Agroforestry practices increased tree biomass and reduced soil

erosion (correlation coefficient = -0.31). The increase in tree biomass also meant that households received energy provisioning ecosystem services through increased availability of firewood. This effect was especially important for women, who are typically responsible for collecting firewood. Agroforestry led to significant reductions in fuel wood purchased (49 percentage point reduction in households purchasing firewood in Lower Nyando and a 4% point reduction in households purchasing firewood in Middle Nyando) and the time spent collecting firewood (a 180-minute average reduction in Lower Nyando and 40-minute average reduction in Middle Nyando).

From our agroforestry EGM, we found that studies of agroforestry practices (without a specific intervention) primarily focused on ecosystem service outcomes, such as biodiversity conservation and soil and water quality. These outcomes are well-studied in terms of practices but not in the context of specific interventions. Ecosystem service outcomes are relatively understudied in the impact evaluations, and there is no consistency in environmental indicator variables used. However, what has been studied indicates that the environmental benefits are being achieved to at least some extent, consistent with the broader literature on agroforestry practices.

Intersection of agroforestry for environmental and social outcomes

In much of the evidence on agroforestry interventions, agroforestry is assumed to have underlying environmental benefits, but researchers rarely evaluate both outcomes within the same study. The baseline assumptions in many of the studies are that agroforestry provides one or more of the following ecosystem services: protect crops from wind and soil erosion, help farmers in climate change mitigation and adaptation, improve soil fertility, sustain biodiversity, and sequester carbon. Most of the impact evaluations, however, did not evaluate the impacts of the programs on these outcomes, instead focusing on food security and economic outcomes. Only four studies included in our review examined both social and environmental outcomes (Hagggar et al., 2017; Pender et al., 2008; Sills & Caviglia-Harris, 2015; Thorlakson & Neufeldt, 2012), with four others looking at both agricultural yield and income outcomes (Hegde & Bull, 2011; Kuntashula & Mungatana, 2013; Place et al., 2005, Thorlakson & Neufeldt, 2012). While we had anticipated exploring the interaction between environmental and social outcomes, the evidence base was too minimal to evaluate the intersection of these outcomes.

The role of interventions in the adoption and success of agroforestry

Extant evidence suggests two distinct pathways through which agroforestry interventions work. The first is *provision of information, raw materials, and technical support*, which includes support for activities like extension and training programs and germplasm provision. Such interventions work under the assumption that the benefits of the agroforestry are additive, with environmental benefits supporting additional social and economic benefits. Related practices primarily include agroforestry for soil fertility replenishment, such as improved fallow and fertilizer trees. Under this scenario, when agroforestry interventions are introduced and individuals adopt

agroforestry practices, the practices lead to improved yields and income generation as well as ecosystem service generation.

Under this pathway, three key elements, participant engagement, program exposure, and indirect financial support, lead to agroforestry adoption, which is then intended to lead to positive productivity, profitability, and human well-being outcomes through diversified production and income streams, improved soil fertility, and other changes. The promoted practices are simultaneously supposed to lead to improved ecosystem services outcomes, such as soil and water management, soil fertility replenishment, habitat provision, and carbon sequestration.

Bostedt et al. (2016), Coulibaly et al. (2017), Kuntashula and Mungatana (2013), Thorlakson and Neufeldt (2012), Dai et al. (2017), Sills and Caviglia-Harris (2015), and Place et al. (2005) examined interventions along this pathway. The studied interventions tended to lead to increased food security through improved yields or dietary diversity. Coulibaly et al. (2017) explicitly analyzed the impact of the agroforestry program on adoption of agroforestry practices, and they found that receiving agroforestry training significantly increased the likelihood of adopting agroforestry practices through building farmer capacity to appropriately manage the trees. Place et al. (2005) found that 22% of farmers in the pilot villages adopted fertilizer trees, from about 10% of farmers who were practicing agroforestry at the start of the program. The importance of longer-term studies of the impacts of agroforestry interventions and of appropriate intervention targeting for these types of interventions was also highlighted. Mixed results were found in several studies, with one study showing a negative impact of the program on household income (Dai et al., 2017), and there were changes in effects over different time scales.

The second type of intervention pathway we identified is *incentive provision* to enhance value and offset economic tradeoffs. Interventions following this pathway include a range of incentive provision approaches such as certification programs and PES. Certification programs are intended to increase the prices for sustainably produced products. PES comprise direct payments to farmers for maintaining a specified set of sustainable practices. These interventions rely on the provision of incentives to offset the decreases in productivity or profitability that occur when sustainable practices are adopted. These interventions are useful when the practice involves taking land out of productive use or incorporating elements that decrease the overall productivity of the system. By providing financial incentives, farmers are motivated to adopt the desired practices since the sustainable alternatives maintain the same or higher profitability as the conventional options.

Like the first pathway, these interventions also require participant engagement and program exposure through promotion of the program to achieve adoption. Adoption of the promoted agroforestry practice is specifically intended to achieve environmental benefits, such as soil and water quality, biodiversity conservation, and carbon sequestration. In this case, however, the planted trees take land out of production without any direct increase in productivity or productivity due to the practice. Direct financial incentives are therefore required to offset the loss in productivity. The intervention may be in the form of a payments for environmental services program that directly pays farmers for

implementing and maintaining sustainable management practices or in the form of a certification program that brings in higher prices for products produced under sustainable management practices. When these incentives provide enough additional income to offset or exceed the loss of profits due to decreased productivity, farmers are incentivized to implement the promoted practices. However, even with these direct offsets, an increased labor requirement, increased difficulty in farm management, or lack of market access may still deter farmers from adopting the promoted agroforestry practice(s).

Haggar et al. (2017), Hegde and Bull (2011), Pagiola et al. (2016), and Pender et al. (2008) analyzed interventions within this second pathway. Studied programs incorporated incentives to promote the adoption of agroforestry practices to offset the costs of taking land out of production or reducing the productivity of the system. Two of the projects assessed were pilot projects, one was an early-stage assessment, and one was a later stage assessment of the program. None of the studies explicitly evaluated the rate of participation in the different programs.

Equity focus of agroforestry interventions

We found three studies that explicitly considered variable impacts across different sub-groups of the study population, but most included studies made some suggestions about group differences. The studies considered differential impacts on small-holders versus large-holders, on woman-headed households versus male-headed households, and on richer groups versus poorer groups.

Coulibaly et al. (2017) found that households with the smallest landholdings were most positively impacted by the intervention. This program targeted smallholders and the poorest populations, and they found that this was key to the success of the program—they found a high treatment effect on the treated, but a small, expected treatment effect on the untreated. On the other hand, Kuntashula and Mungatana (2013) predicted that the average treatment effect on the untreated was much higher than the treatment effect on the treated group. Through discussion with farmers, they found that there were several barriers to entry that prevented some households who may have received the highest benefits through the program from entering the program. The main barrier cited by farmers was the long waiting time for the accrual of benefits of planting trees, which require several years to grow before providing benefits.

Female-headed household were often disproportionately affected or overlooked by agroforestry interventions, with some of the studies showing that woman-headed households have less positive or more negative impacts on their households than male-headed households. Even when a program did not benefit men more than women, there were baseline differences between the two groups such that women-headed households in both the treatment and control groups had lower food crop value and income. Hegde and Bull (2011) explicitly considered the differences between male-headed and female-headed households, and they found significantly higher benefits for male-headed households. They use decomposition analysis to find that 54% of the differences were due to different endowments, and 46% was due to discrimination. Place et al. (2005) noted that it was difficult for women to participate in agroforestry training programs because of restrictions from their

husbands and household chores, and they could not receive the benefits of the intervention due to insufficient landholdings. Thorlakson and Neufeldt (2012) noted the disproportional benefit to women of having access to firewood, since women are often responsible for collection and have to walk miles to collect firewood if they are without trees on their farm.

There was also considerable variability in how agroforestry interventions impacted poor or marginalized households. Some projects specifically targeted the poorest households, while others found that poor households had less positive or more negative impacts on their households than richer households. Hegde and Bull (2011) explicitly considered the differences between male-headed and female-headed households, and they found significantly higher benefits for male-headed households. They use decomposition analysis to find that 36% of the differences were due to different endowments, and 64% was due to discrimination. Several of the studies suggested that poor households are expected to receive the highest benefits from agroforestry interventions.

6 | DISCUSSION

6.1 | Summary of main results

We identified 11 impact evaluations of agroforestry interventions across nine different countries. We were able to quantitatively synthesize results for the impacts of agroforestry interventions on crop yield and income through meta-analysis techniques. We additionally used narrative synthesis for impacts of agroforestry interventions on nutrition and human-wellbeing and environmental outcomes.

Our meta-analysis showed the average effect of these interventions on crop yield outcomes was a large size and positive (1.16 standardized mean difference (Hedges' g) with a 95% CI of $[-0.35, 2.67]$) but was not statistically significant ($p = .13$). There was substantial variability in the results for crop yield. The average effect of interventions in our meta-analysis of household income outcomes was very small but positive (0.12 standardized mean difference (Hedges' g) with a 95% CI of $[-0.06, 0.30]$), and it was not statistically significant ($p = .20$). Results were less variable than for yields, though there was still moderately high heterogeneity.

There was a positive or neutral impact of agroforestry on nutrition and food security from the studies that considered one or both as an outcome measure. However, there was not enough evidence to perform meta-analysis. This lack of evidence is a major gap in agroforestry research, as agroforestry is often promoted for diversifying food production. Although the evidence base is thin, it currently favors the hypothesis that agroforestry interventions can improve nutritional and food security outcomes.

There was also a notable lack of evidence on environmental outcomes. Most of the studies assumed that the environmental benefits came with the implementation of agroforestry practices based on previous work studying the practices alone. While the environmental impacts of agroforestry practices are relatively

well-studied (Miller et al., 2020), there is a lack of evidence that these benefits translate in the context of a specific intervention.

6.2 | Overall completeness and applicability of evidence

There was a considerable overall gap in evidence available on the impacts of agroforestry interventions. We only identified 11 impact evaluations of agroforestry interventions, and the 11 lacked consistency in the measurement of indicator variables and analytical techniques. The available impact evaluations studied a wide range of practices and types of interventions, making it difficult to compare across different contexts and situations. However, the results do provide a baseline to inform future research, and the results revealed trends in agroforestry intervention pathways.

Agroforestry is often promoted for its promise to deliver on both environmental and social outcomes, but evidence supporting this claim in the context of specific interventions remains lacking. Few studies examined cobenefits and tradeoffs of agroforestry in the context of interventions. Agrawal and Chhatre (2011) argued for the importance of considering multiple social and ecological outcomes at the same time in coupled natural and human systems. Our results suggest that their call has been heeded more rarely than required to advance integrated understanding of the multiple outcomes of agroforestry interventions. There remains an urgent need to simultaneously examine multiple outcomes of agroforestry interventions to inform efforts to achieve the 2030 UN Sustainable Development Goals.

6.3 | Quality of the evidence

The overall quality of the evidence base was low, with all of the studies rating as “critical” (55%) or “high” (45%) for overall risk of bias. The biggest issue was that all of the studies were based on interventions with self-selection into the program and only used cross-sectional data at the end line in the analysis. All of the studies used methods to try to address this selection bias through PSM, endogenous switching regression, Heckman two-stage regression, difference-in-difference, and IV techniques. The studies that used matching mostly matched on a limited number of variables and did not account for several important factors that could affect the selection process and outcomes. This approach also led to issues with group equivalence. We also identified several issues with potential spillover effects that were not controlled for in the analysis. While the overall risk of bias was high, these studies represent an important advance as the field of agroforestry seeks to better understand program effectiveness.

6.4 | Potential biases in the review process

We tried to limit the potential bias within the review process by double screening of studies for inclusion and double extracting data

where possible. The risk of bias assessment was also performed by two separate reviewers and discussed until reviewers agreed on a rating to assign. The effect size extraction and calculation were done by one lead reviewer and checked by another lead reviewer. When a paper did not provide enough information to calculate effect size, the study authors were contacted, and the additional information was retrieved.

The review only includes studies through October 2017 and those that were published in English, which is a limitation of our review. A number of relevant impact evaluations have been published since we conducted our search. We intend to update our SR with new literature published since 2017 within 5 years of our previous search through the year of our updated publication.

Due to resource constraints, we only included studies in English, which can limit the scope of our results by not including relevant studies in other languages. For example, our study has likely missed important evidence described in French, Mandarin Chinese, Portuguese, and Spanish, among other languages. Similarly, given that the practice of agroforestry has different names in different places, is possible that we missed a relevant term in our search strategy, even though the terms we used were developed in consultation with a search specialist and our advisory team, which included several experts in this field. There may also be geographical biases since we included only English language studies.

As previously discussed, we decided to exclude adoption-only impact evaluations, which could offer a useful counterpart to this review. While the literature on agroforestry practices is extensive, it also indicates mixed results of implementing agroforestry practices. It would therefore be difficult to estimate even the direction of the social and environmental impacts of agroforestry adoption.

The targeted scope of this SR and the search limitations may contribute to small size of the evidence base; however, we believe that we captured much, if not all, of the literature evaluating the impacts of agroforestry interventions within our timeframe. The scope of our EGM was very broad, and presentations to relevant stakeholders and our advisory committee revealed no additional known studies relevant for inclusion.

6.5 | Agreements and disagreements with other studies or reviews

While there are no similar reviews of agroforestry interventions, previous reviews have synthesized evidence on the effectiveness of certification schemes and PES programs, which were policy instruments used for several agroforestry interventions included in our review (Oya et al., 2017; Snilstveit et al., 2019). Oya et al. (2017) reviewed 43 quantitative and 136 qualitative studies and found that certification schemes tend to increase prices and income from produce but had no consistent effect on wages or household income. Like the evidence for agroforestry certification schemes, they found mixed effects on yields but generally positive impacts to income from production (though the effects on total household income were unclear). As with our review, they also concluded that context-specific intervention design is key for successful implementation of certification schemes. Snilstveit et al. (2019) examined

the impacts of PES on environmental and socioeconomic outcomes and identified 44 quantitative impact evaluations and 60 qualitative studies. They two PES studies we included were also included by Snilstveit et al. (2019). They found the PES programs may increase household income and improve environmental outcomes across all PES programs included, which held true when we considered only agroforestry PES programs.

We note that available evidence for the above reviews was much more extensive than was the case for our review: they each included at least five times the number of relevant quantitative studies as the present review.

Several SRs synthesized the evidence on the impacts of agricultural interventions on nutrition outcomes and found that nutrition education and homegardens had the greatest likelihood of inducing positive nutritional outcomes, especially consumption of nutrient-rich crops and dietary diversity, though the results were mixed (Berti et al., 2004; Bird et al., 2019; Masset et al., 2012; Pandey et al., 2016). These SRs also highlighted the importance of women's participation in agricultural interventions to achieve nutritional outcomes. With the agroforestry interventions we identified, the intervention that focused on education similarly led to positive nutritional outcomes (Bostedt et al., 2016), while agroforestry technologies that may provide access to increased food security did not necessarily translate to more diversified nutrition (Place et al., 2005) or did not explicitly measure nutritional status as part of the study (Coulibaly et al., 2017; Kuntashula & Mungatana, 2013).

While we only identified one tenure reform intervention associated with agroforestry practices, which is insufficient for drawing broader conclusions, Lawry et al. (2014) reviewed the impacts of tenure interventions on investment and agricultural productivity generally, which provides more general understanding of tenure interventions. The evidence base for land property rights interventions they examined was limited ($n = 20$), but their findings suggest that land tenure interventions are plausible pathways to improve welfare.

Loevinsohn et al. (2013) suggested in their theory of change similar pathways of change for agricultural technology adoption generally as we found specifically for agroforestry. They also found surprisingly few high-quality studies ($n = 5$) that evaluated the circumstances and conditions where agricultural technology adoption led to changes in agricultural productivity. Finally, Ingram et al. (2016) also highlighted the importance of gender equity in forests, trees, and agroforestry intervention, and they found differences in the nature of the products and activities that men and women participate in as well as evidence of gender differentiated incomes. They also conclude that interventions need to be explicitly gender sensitive, support collective action, and consider parallel actions to reduce gender disparity in forest, trees, and agroforestry interventions.

7 | AUTHORS' CONCLUSIONS

Agroforestry has been widely practiced, promoted, and studied across the LMICs of Africa, Asia, and Latin America. Given its prevalence and promise, agroforestry is promoted for its potential to provide a vital contribution to advancing several of the 2030 UN SDGs

(Van Noordwijk et al., 2018; Waldron et al., 2017). Indeed, high-level policy documents in many LMICs now explicitly call for the integration of trees into farming systems (FAO, 2013) and international donors have invested billions of dollars in agroforestry interventions around the world (AidData, 2017; Tierney et al., 2011).

In this study we have presented the findings of a SR that used systematic methods to identify, collect, and synthesize available evidence on the effects of agroforestry interventions in LMICs on three important outcomes: agricultural productivity, ecosystem services, and human well-being. Based on the available evidence, we reviewed the impacts of specific agroforestry interventions on crop yields, income, nutrition and human well-being, and environmental outcomes. The main finding of our review is that there is a critical lack of rigorous evidence on the impacts of agroforestry interventions. However, the existing evidence points to the positive or neutral impacts of agroforestry interventions on multiple social and ecological outcomes.

7.1 | Implications for practice

Our review offers a baseline of the impacts of agroforestry interventions to guide future practice and policy. Given that the major finding is that there is a critical gap in evidence on the effects of the agroforestry interventions, it may seem there are few implications for policy and practice. In a sense this is true—we lack systematic understanding of the relative effectiveness of different interventions to inform new policies and programs. However, the overall findings of this report do suggest some important paths forward.

First, the review highlights the need for additional funding for impact evaluations of agroforestry programs and policies. There is a need for donors to explicitly call for rigorous impact evaluations as part of the implementation of the interventions. While in many cases the evidence suggests there may be positive impacts of agroforestry interventions, the evidence base is extremely limited. We suggest that careful piloting and baseline assessment occur as a prerequisite of new program implementation in new contexts. Future studies should carefully consider using RCT designs to test the effectiveness of different intervention approaches. Research in agriculture (e.g., Carter et al., 2013; Jack & Cardona Santos, 2017; Weiser et al., 2015) and forestry (e.g., Jayachandran et al., 2017) as well as a raft of agroforestry field trials (Miller et al., 2020) suggests the feasibility of RCTs in this domain.

Second, the setting and the type of agroforestry practice promoted are crucial considerations for intervention targeting and effectiveness. For example, in areas with low productivity and soil fertility and high poverty and food insecurity, interventions may prioritize income and yield outcomes by promoting fertilizer trees or improved fallow rather than shade trees or trees for carbon sequestration. These interventions may simultaneously target poverty reduction and food security, with less emphasis on biodiversity conservation and carbon sequestration. Evidence reviewed here suggests that such interventions may be suitable for smallholder farmers in highly degraded areas, though additional studies are necessary to test this hypothesis given the small number of studies. In areas with high biodiversity or high potential to sequester carbon,

the focus may turn to environmental outcomes. Practices associated with environmental outcomes, such as shade trees in coffee plantations or pastures or intercropped trees for carbon sequestration, potentially reduce yields, which may require income offsets such as higher prices through certification programs or PES. Ideally, under either scenario, incomes will be increased to promote the adoption of the practices, benefit the landholders, and help improve human well-being. In many cases, agroforestry systems can incorporate more diverse food sources, such as fruit and nut trees, to improve nutrition outcomes as well.

Finally, our results show that the impacts of agroforestry interventions were experienced differently by different population subgroups. New programs should therefore consider who and how they are targeting groups in program implementation, with special attention to gender and social class. We found that female-headed households were often disproportionately affected or overlooked in the studied agroforestry interventions, with some results showing that woman-headed households experienced less positive or more negative impacts than male-headed households. There was also considerable variability in how agroforestry interventions impacted poor or marginalized households. Some studied projects specifically targeted the poorest households. Results were mixed, with some finding that poor households received the most benefit, while others found that such households had less positive or more negative impacts on their households than wealthier households. Several studies reported variable differences between the average treatment effects on the treated versus the average treatment effects on the untreated. These variations highlight the need to carefully consider the targets of an intervention and understand the incentives and barriers to entry.

7.2 | Implications for research

Our study reveals that rigorous evidence on the effects of agroforestry interventions remains extremely limited. Impact evaluations of agroforestry interventions remain challenging due to the long timescale between implementation and impacts. Trees take a long time to grow, and the resulting effects on environmental health and human livelihoods may take decades. The scope of many development projects usually only lasts a few years, so long-term monitoring and evaluation must be built into project proposals and designs. Many studies we found only examined whether farmers adopted agroforestry as the results of an intervention, without measuring the subsequent impacts on social-ecological outcomes, so they were not included in our review. One approach to addressing the need for long-term evaluation is establishing on-farm experimental trials with treatment and control farmers, for which there may be better justification for long-term monitoring proposals. Finally, RCTs are rarely conducted in agroforestry research based on our findings, but RCTs can offer valuable insights into how agroforestry interventions impact farmer livelihoods and the environment.

The complexity that comes with integration of agricultural, forest, and pastoral, and other systems, as done in agroforestry, poses significant challenges to evaluating the effectiveness of specific agroforestry interventions. However, given the potential of agroforestry to

contribute to a number of major sustainable development goals simultaneously, there is an urgent need for such impact evaluation. Nevertheless, there are examples demonstrating such evaluation is possible. Expanding the number of impact evaluations of agroforestry interventions, especially using RCTs, therefore represents a major opportunity for expanding and improving the existing evidence base. There is also a need for long-term trials with baseline data collected to conduct impact evaluations using panel data to understand how the impacts of the agroforestry interventions change over time.

A better understanding of the win-win scenarios and tradeoffs associated with agroforestry is urgently needed, particularly given the potential of agroforestry to help achieve the SDGs. More robust evidence on the different environment and development objectives agroforestry can advance, including climate change mitigation and adaptation, poverty reduction, and health and nutrition, is needed in its own right, but also to enable analysis of synergies and tradeoffs.

Agroforestry encompass a huge suite of different practices that are flexible in their design and composition. This spectrum of practices that agroforestry captures makes it difficult to define for comparison to alternative land uses. Coupling the ecological suitability of different agroforestry practices with the associated impacts on human well-being, instead of leaving the ecology and human well-being outcomes separate, could build understanding of the complex dynamics of agroforestry to help design better agroforestry interventions. We need to better understand the costs and benefits of agroforestry from an interdisciplinary perspective, incorporating economics, social science, and environmental science, to assess its viability as a conservation practice while also considering the needs of farmers.

ACKNOWLEDGEMENTS

We thank members of the advisory group to the agroforestry evidence gap map (EGM), Kahlil Baker, Samantha Cheng, Steve Franzel, Irina Klytchnikova, and Frank Place. Special thanks to Birte Snilstveit for supporting, guiding and peer reviewing this work at each stage in its development. We are especially grateful to Festus Amadu, Delaney Demro, Samantha Irwin, Protensia Hadunka, Rebecca Laurent, Nataly Medina, Nikolas Merten, Joseph Murray, Katia Nakamura, Noé Nava, Pushpendra Rana, Krti Tallam, Kealie Vogel, and Adam Yusen for research assistance on this project. Nikolas Merten assisted with the systematic review data extraction and as an additional coder for risk of bias analysis. We thank three anonymous reviewers and the editor of Campbell International Development Coordinating Group, Dr. Maren Duvendack, for their feedback on an earlier version of the manuscript, which greatly improved the quality of this report.

CONTRIBUTIONS OF AUTHORS

- Content: Sarah E. Castle wrote the first draft of the manuscript and led the revision in response to the reviewers. All other authors contributed some text and reviewed the manuscript at each stage of its production. Systematic review methods: Sarah E. Castle, Daniel C. Miller, Pablo J. Ordonez, and Kathy Baylis have substantial

understanding of SR methods. The team also consulted with Birte Snilstveit, a leading expert in SR methods, throughout the SR process.

- **Statistical analysis:** Sarah E. Castle, Daniel C. Miller, Pablo J. Ordonez, and Kathy Baylis have expertise in econometrics and statistics. Sarah E. Castle created the R code for the analysis and figures.
- **Information retrieval:** Sarah E. Castle and Pablo J. Ordonez led the information retrieval and data management. The team consulted an advisory group of experts in agroforestry and systematic searches to help create the search string. The team also consulted library scientists at the University of Illinois at Urbana-Champaign for guidance in searching multiple academic databases and grey literature sources.

DECLARATIONS OF INTEREST

The authors declare that there are no conflicts of interests. None of the authors are researchers on any of the studies examined in this review. All authors have published work within the field of agroforestry, which may be cited in the background section of this review. Karl Hughes is an employee of ICRAF, an organization that funded and/or implemented many of these interventions.

PLANS FOR UPDATING THIS REVIEW

We plan perform a follow-up review to update this review within 5 years. In the follow-up review, we plan to include studies that evaluate the impact of agroforestry interventions on agroforestry adoption as well as social-ecological outcomes.

DIFFERENCES BETWEEN PROTOCOL AND REVIEW

Our EGM deviated from the protocol in several ways as the review team identified challenges in screening. First, we did not include researcher-managed agroforestry field trials due to time and resource constraints and due to conceptual differences between researcher-managed fields under controlled experimental conditions and fields managed by farmers. Secondly, we did not include studies that used other types of agroforestry practices as a comparator to an agroforestry practice, as opposed to using a nonagroforestry land use as a comparator. After reviewing began, we realized that these studies aimed to optimize agroforestry configurations, rather than demonstrate the impact of agroforestry practices. Thirdly, we limited our time scope to start in the year 2000, rather than 1990, due to time and resource constraints. For our parallel systematic map for high-income countries (Brown et al., 2018), we considered studies back to the year 1990, and noticed drastically diminished returns prior to 2000 due to limited evidence meeting our inclusion criteria (e.g., many studies did not include a relevant comparator or relevant outcome, rather the focus was more on tree species selection and breeding).

These decisions did not impact our systematic review. The earliest study we found using quasi-experimental impact evaluation methods was the year 2005, and we do not believe there were any published prior to the year 2000 since these methods are relatively new for the fields of agriculture and conservation. We found no evidence of impact evaluations comparing different types of agroforestry as the only comparator. Researcher-managed field trial studies would not have been included in this systematic review of agroforestry interventions.

However, if these field trials were a component of an intervention and the impact evaluation assessed the impacts of the intervention on farmers, we would have included them.

CHARACTERISTICS OF STUDIES

CHARACTERISTICS OF INCLUDED STUDIES

Bostedt et al. (2016)

Methods	Heckman two-stage probit regression
Participants	291 farmers (164 received advice, 127 did not)
Interventions	Farmer capacity development Enhancing access to tree germplasm
Outcomes	Food security and nutrition
Notes	

Risk of Bias

Bias	Authors' judgement	Support for judgement
1: Mechanism of assignment: was the allocation or identification mechanism able to control for selection bias?	High risk	Issue of self-selection to receive assistance from Vi Agroforestry as farmers chose whether to take advantage of their interventions
2: Group equivalence: was the method of analysis executed adequately to ensure comparability of groups throughout the study and prevent confounding?	High risk	Although this study uses a Heckman 2-stage estimation procedure to account for bias, there is self-selection and baseline data was not used in this analysis
3: Performance bias: was the process of being observed free from motivation bias?	Low risk	Individuals randomly selected for survey and uniformly surveyed across treatment and control groups. The authors do not discuss how the monitoring and interviews could adjust the performance. There does not appear to be a risk of the Hawthorne or John Henry effect; They randomly sampled individuals to survey among treatment and control groups

4. Spill-overs, cross-overs and contamination: was the study adequately protected against spill-overs, cross-overs and contamination?	High risk	Self-selection to receive Vi Agroforestry advice (the treatment), no barrier to prevent those who received advice from sharing with those who did not receive advice; not addressed in paper; authors state that there could have been information spread from those that received assistance and information from Vi Agroforestry to those who had not formally been involved with the program
5. Outcome measurement bias	Low risk	Treatment and control groups received same survey at same time without apparent incentive give biased responses; not addressed in paper; There was not an "independent party" conducting the survey. Though, some monitoring took place to confirm information about agroforestry obtained from interviews
6. Selective analysis reporting: was the study free from selective analysis reporting?	Unclear risk	All relevant outcomes reported; Significance reported from probit model from first stage selection in Heckman model. Not all of the requirements for tests in d) were included
7. Other risks of bias: Is the study free from other sources of bias?	Low risk	Reported results do not suggest any other sources of bias

Coulibaly et al. (2017)

Methods	Endogenous switching regression Instrumental variables
Participants	338 farmers (135 adopters, 203 nonadopters)
Interventions	Farmer capacity development Enhancing access to tree germplasm
Outcomes	Productivity (yield) Profitability Income and household expenditure Food security and nutrition
Notes	

Bias	Authors' judgement	Support for judgement
1: Mechanism of assignment: was the allocation or identification mechanism able to control for selection bias?	Unclear risk	Self-selection to participate in training program. The instruments they used were participation in agroforestry training, which they argued affected the decision to adopt but did not affect the dependent variable except through adoption. They tested this instrument and found that it was positively correlated with the decision to adopt agroforestry but was not significantly related to food productivity. However, it is likely that there are systematic differences between farmers who decide to participate in agroforestry training programs and those who do not
2: Group equivalence: was the method of analysis executed adequately to ensure comparability of groups throughout the study and prevent confounding?	Unclear risk	Hausman test not reported; insufficient details provided. Appropriate statistical tests not reported; exogeneity of instrument in question
3. Performance bias: was the process of being observed free from motivation bias?	Low risk	Survey conducted by well-trained enumerators; measured adopters and nonadopters from both program participant group and nonparticipant group; and a pre-test was conducted ahead of time
4. Spill-overs, cross-overs and contamination: was the study adequately protected against spill-overs, cross-overs and contamination?	Low risk	Clearly defined rules for adoption, all others nonadopters. The study dedicates an entire section to how they distinguish participants and nonparticipants: The study has well defined adopters and nonadopters that would prevent spill-over bias,

		etc. There is no obvious problem but there is no information reported on potential risks related to spillovers, contamination, or survey effects in the control group
5. Outcome measurement bias	Low risk	There is a potential for recall bias, but it is likely same between groups; timing same between groups
6. Selective analysis reporting: was the study free from selective analysis reporting?	Unclear risk	Relevant outcomes reported, but appropriate tests not reported
7. Other risks of bias: Is the study free from other sources of bias?	Low risk	Reported results do not suggest any other sources of bias

Dai et al. (2017)

Methods	Propensity score matching
Participants	352 households (236 participants, 116 nonparticipants)
Interventions	Farmer capacity development Enhancing access to tree germplasm
Outcomes	Income and household expenditure
Notes	

Risk of Bias

Bias	Authors' judgement	Support for judgement
1: Mechanism of assignment: was the allocation or identification mechanism able to control for selection bias?	High risk	Self-selection to participate in program
2: Group equivalence: was the method of analysis	High risk	Self-selection to participate in program

		executed adequately to ensure comparability of groups throughout the study and prevent confounding?
3. Performance bias: was the process of being observed free from motivation bias?	Low risk	Consistent survey across treatment and control conducted at same time. Households sampled randomly
4. Spill-overs, cross-overs and contamination: was the study adequately protected against spill-overs, cross-overs and contamination?	Unclear risk	Not discussed by authors
5. Outcome measurement bias	Unclear risk	Not discussed by authors
6. Selective analysis reporting: was the study free from selective analysis reporting?	Low risk	Relevant outcomes reported, some appropriate tests reported
7. Other risks of bias: Is the study free from other sources of bias?	Low risk	Reported results do not suggest any other sources of bias

Haggar et al. (2017)

Methods	Propensity score matching
Participants	278 farms (202 certified farms (see below), 76 noncertified farms)
Interventions	Incentive provision
Outcomes	Reg & Main—Physical, chemical, biological conditions Reg & Main—Mediation of flows Reg & Main—Mediation of waste, toxics and other nuisances Income and household expenditure Productivity (yield)
Notes	

Risk of Bias

Bias	Authors' judgement	Support for judgement
1: Mechanism of assignment: was the allocation or identification mechanism able to control for selection bias?	High risk	Self-selection into certification program; only statistical propensity score matching done
2: Group equivalence: was the method of analysis executed adequately to ensure comparability of groups throughout the study and prevent confounding?	High risk	The certified and noncertified units are clear; there is a matching based on characteristics that would not have been changed by certification and they appear to be relevant. Only matched on three variables: area under coffee, altitude, and level of education of farmer. No baseline data available to match the participants or groups on
3. Performance bias: was the process of being observed free from motivation bias?	Low risk	Consistent survey across treatment and control conducted at same time. The survey was conducted and surveyors also confirmed the information provided. Participation in certification appears to be independent of the monitoring process. However, the authors do not mention the potential risks that the monitoring could have on the outcomes measured
4. Spill-overs, cross-overs and contamination: was the study adequately protected against spill-overs, cross-overs and contamination?	Unclear risk	Self-selection to certify farm, uncertified farms unlikely to incorporate certification practices unless they are in the process of certification; not discussed in paper. There is unlikely to be spill-over

because there is a clear certification threshold in this study. This creates two distinct groups to measure. Though, the authors selected certified and noncertified farms within close proximity; and although a farm may not have been certified, it could have mimicked the practices used by certified farms. This is unlikely because there would not be much benefit for a farmer to adhere to certification standards without being certified, but this aspect was not specifically addressed by the authors (authors do not discuss the chance of participant/nonparticipants being in a transition stage where they have adopted practice, but are not yet certified)

5. Outcome measurement bias	Low risk	Treatment and control groups received same survey at same time without incentive give biased responses; not addressed in paper. Outcome measures are not likely to be biased by the participants. There is also not a benefit for responders to over or under report here. Measurements were made by surveyors to verify the data being collected
6. Selective analysis reporting: was the study free from selective analysis reporting?	Unclear risk	There is no evidence to suggest that the outcomes of the study were selectively reported in the results. Both significant and not-significant results

		were shared. Relevant outcomes reported, but appropriate tests not reported
7. Other risks of bias: Is the study free from other sources of bias?	Low risk	Reported results do not suggest any other sources of bias

Hegde and Bull (2011)

Methods	Propensity score matching
Participants	290 households (105 PES, 185 non-PES)
Interventions	Incentive provision
Outcomes	Productivity (yield) Housing and material assets Income and household expenditure
Notes	

Risk of Bias

Bias	Authors' judgement	Support for judgement
1: Mechanism of assignment: was the allocation or identification mechanism able to control for selection bias?	High risk	Self-selection into PES program; only statistical propensity score matching done
2: Group equivalence: was the method of analysis executed adequately to ensure comparability of groups throughout the study and prevent confounding?	High risk	Self-selection into program; No baseline data available to match the participants or groups on
3. Performance bias: was the process of being observed free from motivation bias?	Low risk	Consistent survey across treatment and control conducted at same time
4. Spill-overs, cross-overs and contamination: was the study adequately protected against spill-overs, cross-overs and contamination?	Unclear risk	Self-selection to join PES, nonparticipants may be likely to incorporate practices if they want to join PES program; not discussed in paper

5. Outcome measurement bias	Low risk	Treatment and control groups received same survey at same time without incentive give biased responses; not addressed in paper
6. Selective analysis reporting: was the study free from selective analysis reporting?	High risk	Relevant outcomes reported, but appropriate tests not reported
7. Other risks of bias: Is the study free from other sources of bias?	Low risk	Reported results do not suggest any other sources of bias

Kuntashula and Mungatana (2013)

Methods	Propensity score matching Endogenous switching regression
Participants	324 households (113 adopters, 213 nonadopters)
Interventions	Farmer capacity development
Outcomes	Productivity (yield) Profitability Income and household expenditure Food security and nutrition
Notes	

Risk of Bias

Bias	Authors' judgement	Support for judgement
1: Mechanism of assignment: was the allocation or identification mechanism able to control for selection bias?	Unclear risk	Self-selection to participate; endline + recall data only; used endogenous switching regression to address selection bias
2: Group equivalence: was the method of analysis executed adequately to ensure comparability of groups throughout the	High risk	Self-selection to participate; endline with recall data only

study and prevent confounding?		
3. Performance bias: was the process of being observed free from motivation bias?	Low risk	Consistent survey across treatment and control conducted at same time. The study uses a survey that evaluations the impact of an intervention well after adoption. Monitoring does not appear to have any impact on the performance of participants given the timeframe of adoption and the presence of monitoring/surveying
4. Spill-overs, cross-overs and contamination: was the study adequately protected against spill-overs, cross-overs and contamination?	Low risk	Clearly defined rules for adoption, all others nonadopters; not addressed in paper
5. Outcome measurement bias	Low risk	Treatment and control groups received same survey at same time without incentive give biased responses; not addressed in paper. The study appears to satisfy the above conditions and the participants do not have an incentive to misreport data
6. Selective analysis reporting: was the study free from selective analysis reporting?	Low risk	All relevant outcomes reported; multiple methods tested. There was a pre-test conducted, no evidence that outcomes were selectively reported, and Kernel matching was employed
7. Other risks of bias: Is the study free from other sources of bias?	Low risk	Reported results do not suggest any other sources of bias

Pagiola et al. (2016)

Methods	Difference-in-differences
Participants	98 households (69 PES, 29 non-PES)

Interventions	Incentive provision Farmer capacity development
Outcomes	Reg & Main—Physical, chemical, biological conditions
Notes	

Risk of Bias

Bias	Authors' judgement	Support for judgement
1: Mechanism of assignment: was the allocation or identification mechanism able to control for selection bias?	Unclear risk	Not clear whether mechanism controlled for selection bias
2: Group equivalence: was the method of analysis executed adequately to ensure comparability of groups throughout the study and prevent confounding?	Unclear risk	Insufficient details are provided
3. Performance bias: was the process of being observed free from motivation bias?	Low risk	Control group selected as group who applied to be part of PES group but were not within the first 80 applicants
4. Spill-overs, cross-overs and contamination: was the study adequately protected against spill-overs, cross-overs and contamination?	Unclear risk	No discussion of issue
5. Outcome measurement bias	Low risk	No discussion but little potential risk of participant reporting bias
6. Selective analysis reporting: was the study free from selective analysis reporting?	Low risk	No discussion of issue, but no apparent issue with selective analysis reporting
7. Other risks of bias: Is the study free from other sources of bias?	Low risk	Reported results do not suggest any other sources of bias

Pender et al. (2008)

Methods	Propensity score matching Instrumental Variables
Participants	640 plots (160 HKm, 160 HKm pending, 160 without HKm, 160 park/private)
Interventions	Institutional and policy change
Outcomes	Housing and material assets Income and household expenditure
Notes	

Risk of Bias

Bias	Authors' judgement	Support for judgement
1: Mechanism of assignment: was the allocation or identification mechanism able to control for selection bias?	High risk	Self-selection into program; likely selection bias (p. 36)
2: Group equivalence: was the method of analysis executed adequately to ensure comparability of groups throughout the study and prevent confounding?	High risk	Program was self-selected; there is not clear information on matching characteristics. No baseline data
3. Performance bias: was the process of being observed free from motivation bias?	Unclear risk	Survey conducted at same time for both treatment and control groups. Issue not discussed by authors. People that were not in permit program, but applied, were then assigned to a control —this could have impacted their performance and contaminated the control measures
4. Spill-overs, cross-overs and contamination: was the study adequately protected against spill-overs, cross-overs and contamination?	Unclear risk	Incorporated into analysis by looking at permit holders, those with permit applications, and non-HKm participants, but nonparticipants may have incorporated practices if

		interested in program; not discussed by authors
5. Outcome measurement bias	Unclear risk	Surveys conducted at same time, may be incentive for respondents to give biased responses due to program requirements not being met; Unclear because depending on stage of the permit process or whether a participant group was allowed to get a permit could influence their reporting here
6. Selective analysis reporting: was the study free from selective analysis reporting?	Unclear risk	All outcomes discussed were reported, nonsignificant results reported; unclear methods
7. Other risks of bias: Is the study free from other sources of bias?	Low risk	Reported results do not suggest any other sources of bias

Place et al. (2005)

Methods	Instrumental Variables
Participants	103 households within pilot villages with adopters and nonadopters
Interventions	Farmer capacity development Community-level campaigning and advocacy
Outcomes	Productivity (yield) Food security and nutrition Income and household expenditure Housing and material assets
Notes	

Risk of Bias

Bias	Authors' judgement	Support for judgement
1: Mechanism of assignment: was the allocation or identification mechanism able to control for selection bias?	High risk	Self-selection to participate; endline data only; targeting issue that is causing differences within group, not only self-selection. Differences within groups here based

on how the agents were treating people. For example, inviting some people to meeting and not others even though they were both supposed to receive the treatment

2: Group equivalence: was the method of analysis executed adequately to ensure comparability of groups throughout the study and prevent confounding?	High risk	Appropriate tests not reported
3. Performance bias: was the process of being observed free from motivation bias?	Unclear risk	Consistent survey across treatment and control conducted at same time; not discussed in paper; Intervention and agents monitoring could have caused changes in behavior due to powerful actors within program implementation
4. Spill-overs, cross-overs and contamination: was the study adequately protected against spill-overs, cross-overs and contamination?	High risk	Likely spillover occurred since improved fallow is highly visible and influenced others' decisions on whether or not to be an early tester; scale of issue is not clear. Authors note that there are lots of nearby users which, therefore, could have led to a spill-over issue
5. Outcome measurement bias	Low risk	Treatment and control groups received same survey at same time without incentive give biased responses; not addressed in paper. The implementation of the surveys is unlikely to result in differences in report outcomes from participants

6. Selective analysis reporting: was the study free from selective analysis reporting?	Unclear risk	Not discussed by authors
7. Other risks of bias: Is the study free from other sources of bias?	Unclear risk	Reported that there was favoritism playing in

Sills and Caviglia-Harris (2015)

Methods	Propensity score matching Difference-in-difference
Participants	Varied by year (total sample ranges 195–528 households, 36–64 matched)
Interventions	Farmer capacity development Enhancing access to tree germplasm Market linkage facilitation
Outcomes	Maintenance of physical, chemical, biological conditions Income and household expenditure Housing and material assets
Notes	

Risk of Bias

Bias	Authors' judgement	Support for judgement
1: Mechanism of assignment: was the allocation or identification mechanism able to control for selection bias?	Unclear risk	Self-selection into program, and unclear if all relevant matching variables included. Panel data, but no baseline (no pretest)
2: Group equivalence: was the method of analysis executed adequately to ensure comparability of groups throughout the study and prevent confounding?	High risk	Small sample size with attrition, self-selection into program
3. Performance bias: was the process of being observed free from motivation bias?	Low risk	No obvious issues with monitoring processes. Authors institutions separate from implementing agency

4. Spill-overs, cross-overs and contamination: was the study adequately protected against spill-overs, cross-overs and contamination?	Low risk	Systematic sampling approach reduced the likelihood of spillovers, since immediate neighbors are not included in the sample
5. Outcome measurement bias	Low risk	Treatment and control groups received same survey at same time without incentive give biased responses; not addressed in paper; The study appears to satisfy the above conditions and the participants do not have an incentive to misreport data
6. Selective analysis reporting: was the study free from selective analysis reporting?	Unclear risk	Relevant outcomes reported, but parallel trends assumption not justified clearly
7. Other risks of bias: Is the study free from other sources of bias?	Low risk	Reported results do not suggest any other sources of bias

Thorlakson and Neufeldt (2012)

Methods	Propensity score matching
Participants	116 households (73 participants, 43 nonparticipants)
Interventions	Farmer capacity development Enhancing access to tree germplasm
Outcomes	Productivity Profitability Income and household expenditure Mediation of flows Energy provisioning
Notes	

Risk of Bias

Bias	Authors' judgement	Support for judgement
1: Mechanism of assignment: was the allocation or	High risk	Self-selection to participate in the program; Parameters

identification mechanism able to control for selection bias?		used for matching included: household size, land tenure, household head educational level, soil type, and gender of household head
2: Group equivalence: was the method of analysis executed adequately to ensure comparability of groups throughout the study and prevent confounding?	High risk	No tests reported; insufficient details provided
3. Performance bias: was the process of being observed free from motivation bias?	Low risk	Post-intervention survey, first author institution separate from implementing agency
4. Spill-overs, cross-overs and contamination: was the study adequately protected against spill-overs, cross-overs and contamination?	Unclear risk	No discussion reported by authors
5. Outcome measurement bias	Low risk	Treatment and control groups received same survey at same time without incentive give biased responses; not addressed in paper; The study appears to satisfy the above conditions and the participants do not have an incentive to misreport data
6. Selective analysis reporting: was the study free from selective analysis reporting?	High risk	Relevant outcomes reported, but appropriate tests not reported. Failed to present matching results
7. Other risks of bias: Is the study free from other sources of bias?	Low risk	Reported results do not suggest any other sources of bias

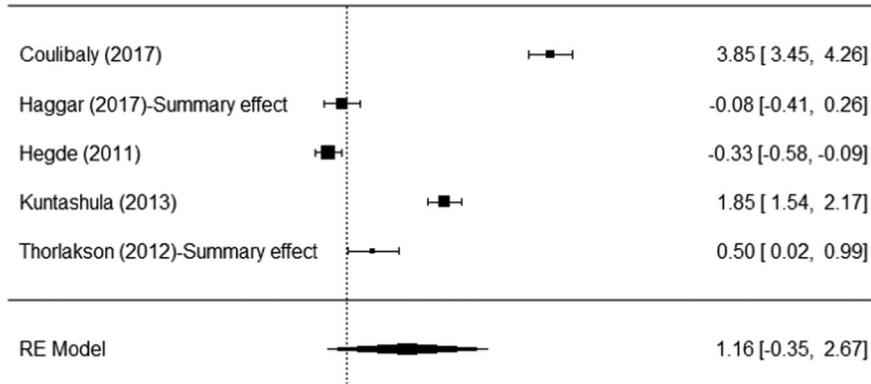
CHARACTERISTICS OF EXCLUDED STUDIES

For the full list of excluded studies, contact the lead author, Sarah E. Castle.

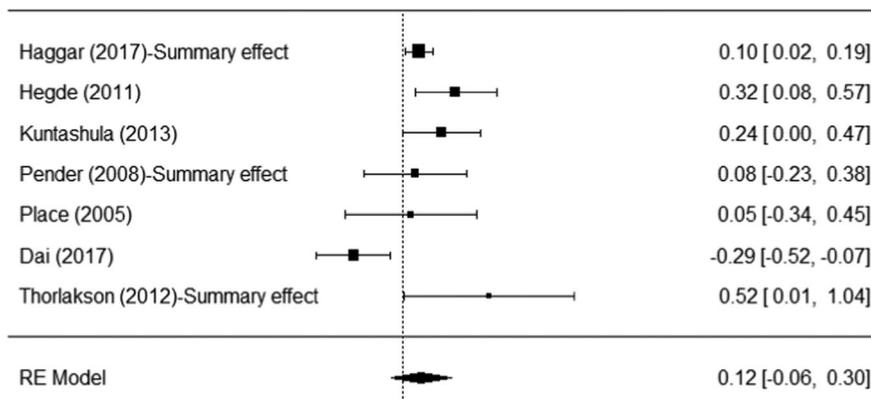
DATA AND ANALYSES

Forest plots

Effects of agroforestry interventions on yields

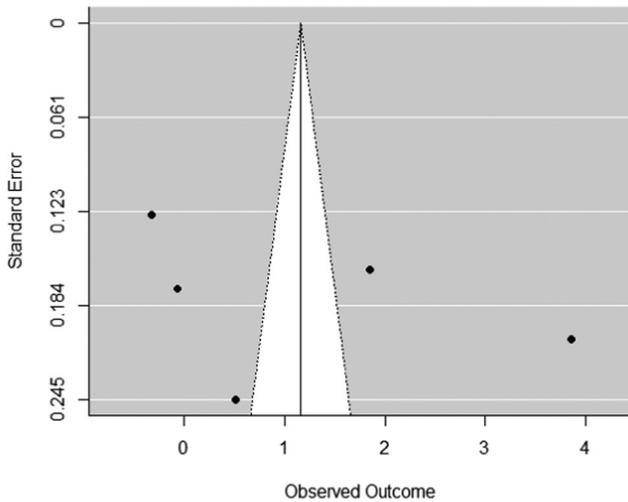


Effects of agroforestry interventions on incomes

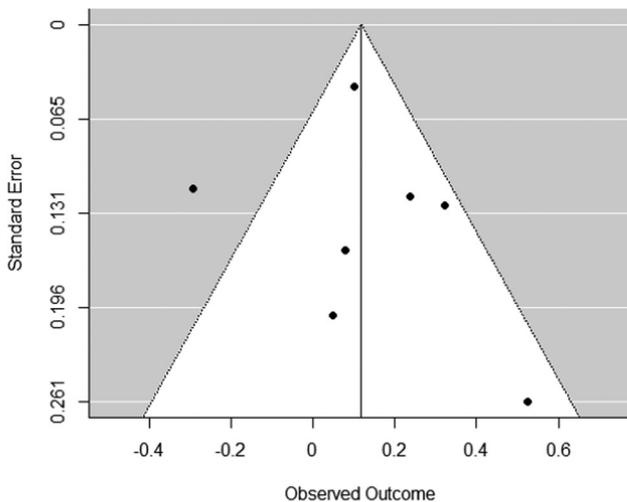


Funnel plots (too few studies to conduct funnel plot analyses)

Effects of agroforestry interventions on yields



Effects of agroforestry interventions on incomes



SOURCES OF SUPPORT

INTERNAL SOURCES

No sources of support provided

EXTERNAL SOURCES

Funding from 3ie and the USDA National Institute of Food and Agriculture, Hatch project #1009327 is gratefully acknowledged.

REFERENCES TO STUDIES

INCLUDED STUDIES

- Bostedt, G., Hörnell, A., & Nyberg, G. (2016). Agroforestry extension and dietary diversity—An analysis of the importance of fruit and vegetable consumption in West Pokot, Kenya. *Food Security*, 8, 271–284.
- Coulbaly, J. Y., Chiputwa, B., Nakelse, T., & Kundhlande, G. (2017). Adoption of agroforestry and the impact on household food security among farmers in Malawi. *Agricultural Systems*, 155, 52–69.
- Dai, X., Pu, L., & Rao, F. (2017). Assessing the effect of a crop-tree intercropping program on smallholders' incomes in rural Xinjiang, China. *Sustainability*, 9, 1542.
- Haggar, J., Soto, G., Casanoves, F., & Virginio, Ed. M. (2017). Environmental-economic benefits and trade-offs on sustainably certified coffee farms. *Ecological Indicators*, 79, 330–337.
- 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.
- Kuntashula, E., & Mungatana, E. (2013). Estimating the causal effect of improved fallows on farmer welfare using robust identification strategies in Chongwe, Zambia. *Agroforestry Systems*, 87, 1229–1246.
- Pagiola, S., Honey-Rosés, J., & Freire-González, J. (2016). Evaluation of the permanence of land use change induced by payments for environmental services in Quindío, Colombia. *PLOS One*, 11, e0147829.
- Pender, J., Suyanto, S., Kerr, J., & Edward, K. (2008). *Impacts of the Hutan Kamasarakatan Social Forestry Program in the Sumberjaya watershed, West Lampung District of Sumatra, Indonesia*. IFPRI Discussion Paper. International Food Policy Research Institute.
- Place, F., Adato, M., Hebinck, P., & Omosa, M. (2005). *The impact of agroforestry-based soil fertility replenishment practices on the poor in western Kenya*. International Food Policy Research Institute.

Sills, E. O., & Caviglia-Harris, J. L. (2015). Evaluating the long-term impacts of promoting “green” agriculture in the Amazon. *Agricultural Economics*, 46, 83–102.

Thorlakson, T., & Neufeldt, H. (2012). Reducing subsistence farmers' vulnerability to climate change: evaluating the potential contributions of agroforestry in western Kenya. *Agriculture & Food Security*, 1, 15.

EXCLUDED STUDIES

List of primary studies (not systematic reviews) included in the EGM (Miller et al., 2020) or identified through the additional search that were excluded for this review. A full list of studies excluded for the EGM will be provided by the authors upon request.

OTHER REFERENCES

Additional references

- Agrawal, A., & Chhatre, A. (2011). Against mono-consequentialism: multiple outcomes and their drivers in social-ecological systems. *Global Environmental Change*, 21, 1–3.
- AidData. 2017. www.aiddata.org
- Amadu, F. O., Miller, D. C., & McNamara, P. E. (2020). Agroforestry as a pathway to agricultural yield impacts in climate-smart agriculture investments: Evidence from southern Malawi. *Ecological Economics*, 167, 106443.
- Atangana, A., Khasa, D., Chang, S., & Degrande, A. (2014). Definitions and Classification of Agroforestry Systems, *Tropical Agroforestry*. Springer Netherlands.
- Bayala, J., Sileshi, G. W., Coe, R., Kalinganire, A., Tchoundjeu, Z., Sinclair, F., & Garrity, D. (2012). Cereal yield response to conservation agriculture practices in drylands of West Africa: A quantitative synthesis. *Journal of Arid Environments*, 78, 13–25.
- Berti, P. R., Krusevec, J., & FitzGerald, S. (2004). A review of the effectiveness of agriculture interventions in improving nutrition outcomes. *Public Health Nutrition*, 7, 599–609.
- Bird, F. A., Pradhan, A., Bhavani, R. V., & Dangour, A. D. (2019). Interventions in agriculture for nutrition outcomes: A systematic review focused on South Asia. *Food Policy*, 82, 39–49.
- Blaser, W. J., Oppong, J., Hart, S. P., Landolt, J., Yeboah, E., & Six, J. (2018). Climate-smart sustainable agriculture in low-to-intermediate shade agroforests. *Nature Sustainability*, 1, 234–239.
- Borenstein, M., Hedges, L., Higgins, J., & Rothstein, H. (2009a). Multiple comparisons within a study, *Introduction to meta-analysis* (pp. 239–242). John Wiley & Sons.
- Borenstein, M., Hedges, L. V., Higgins, J. P., & Rothstein, H. R. (2009b). *Introduction to meta-analysis*. John Wiley & Sons.
- Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2009c). Multiple outcomes or time-points within a study, *Introduction to meta-analysis* (pp. 225–238). John Wiley & Sons.
- Bostedt, G., Hörnell, A., & Nyberg, G. (2016). Agroforestry extension and dietary diversity—An analysis of the importance of fruit and vegetable consumption in West Pokot, Kenya. *Food Security*, 8, 271–284.
- Bottrill, M., Cheng, S., Garside, R., Wongbusarakum, S., Roe, D., Holland, M. B., Edmond, J., & Turner, W. R. (2014). What are the impacts of nature conservation interventions on human well-being: A systematic map protocol. *Environmental Evidence*, 3(16). <http://doi.org/10.1186/2047-2382-3-16>
- Brawn, J. D. (2017). Implications of agricultural development for tropical biodiversity. *Tropical Conservation Science*, 10, 19400829 17720668.
- Brown, S. E., Miller, D. C., Ordóñez, P. J., & Baylis, K. (2018). Evidence for the impacts of agroforestry on agricultural productivity, ecosystem services, and human well-being in high-income countries: a systematic map protocol. *Environmental Evidence*, 7, 24.

- Carter, M. R., Laajaj, R., & Yang, D. (2013). The impact of voucher coupons on the uptake of fertilizer and improved seeds: Evidence from a randomized trial in Mozambique. *American Journal of Agricultural Economics*, *95*, 1345–1351.
- Coe, R., Sinclair, F., & Barrios, E. (2014). Scaling up agroforestry requires research “in” rather than “for” development. *Current Opinion in Environmental Sustainability*, *6*, 73–77.
- Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement*, *20*, 37–46.
- Cohen, J. (1977). Chapter 2—The t test for means. In J. Cohen (Ed.), *Statistical power analysis for the behavioral sciences* (pp. 19–74). Academic Press.
- Corbeels, M., Cardinael, R., Naudin, K., Guibert, H., & Torquebiau, E. (2018). The 4 per 1000 goal and soil carbon storage under agroforestry and conservation agriculture systems in sub-Saharan Africa. *Soil and Tillage Research*.
- Coulbaly, J. Y., Chiputwa, B., Nakelse, T., & Kundhlande, G. (2017). Adoption of agroforestry and the impact on household food security among farmers in Malawi. *Agricultural Systems*, *155*, 52–69.
- De Beenhouwer, M., Aerts, R., & Honnay, O. (2013). A global meta-analysis of the biodiversity and ecosystem service benefits of coffee and cacao agroforestry. *Agriculture, Ecosystems & Environment*, *175*, 1–7.
- Degrande, A., Gyau, A., Foundjem-Tita, D., & Tollens, E. (2014). Improving smallholders' participation in tree product value chains: Experiences from the Congo Basin. *Forests, Trees and Livelihoods*, *23*, 102–115.
- DeLonge, M. S., Miles, A., & Carlisle, L. (2016). Investing in the transition to sustainable agriculture. *Environmental Science & Policy*, *55*, 266–273.
- Ellis, P. D. (2010). Introduction to effect sizes. In P. D. Ellis (Ed.), *The essential guide to effect sizes: Statistical power, meta-analysis, and the interpretation of research results* (pp. 3–30). Cambridge University Press.
- FAO & Buttoud, G. (2013). In F. Place, & M. Gauthier (Eds.), *Advancing agroforestry on the policy agenda: A guide for decision-makers*. Rome: Food and Agriculture Organization of the United Nations.
- Friggens, N. L., Hester, A. J., Mitchell, R. J., Parker, T. C., Subke, J.-A., & Wookey, P. A. (2020). Tree planting in organic soils does not result in net carbon sequestration on decadal timescales. *Global Change Biology*, *26*, 5178–5188.
- Garrity, D. P., Akinnifesi, F. K., Ajayi, O. C., Weldesemayat, S. G., Mowo, J. G., Kalinganire, A., Larwanou, M., & Bayala, J. (2010). Evergreen agriculture: A robust approach to sustainable food security in Africa. *Food Security*, *2*, 197–214.
- Government of India. 2014. National Agroforestry Policy.
- Government of Malawi. 2011. The National Agricultural Policy. Lilongwe, Malawi.
- Haddaway, N. R., Collins, A. M., Coughlin, D., & Kirk, S. (2017). A rapid method to increase transparency and efficiency in web-based searches. *Environmental Evidence*, *6*(1). <http://doi.org/10.1186/s13750-016-0079-2>
- Haggard, J., Soto, G., Casanoves, F., & Virginio, Ed. M. (2017). Environmental-economic benefits and trade-offs on sustainably certified coffee farms. *Ecological Indicators*, *79*, 330–337.
- Hedges, L. V., & Olkin, I. (1985). Chapter 5—Estimation of a single effect size: Parametric and nonparametric methods. In L. V. Hedges & I. Olkin (Eds.), *Statistical methods for meta-analysis* (pp. 75–106). Academic Press.
- 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., Thomas, J., Chandler, J., Cumpston, M., Li, T., Page, M. J., & Welch, V. A. (2020). In Cochrane (Ed.), *Cochrane handbook for systematic reviews of interventions*. John Wiley & Sons.
- Horlings, L. G., & Marsden, T. K. (2011). Towards the real green revolution? Exploring the conceptual dimensions of a new ecological modernisation of agriculture that could 'feed the world'. *Global Environmental Change*, *21*, 441–452.
- Horrigan, L., Lawrence, R. S., & Walker, P. (2002). How sustainable agriculture can address the environmental and human health harms of industrial agriculture. *Environmental Health Perspectives*, *110*, 445–456.
- IAASTD. 2009. Global report international assessment of agriculture at a crossroads.
- Ilstedt, U., Malmer, A., Verbeeten, E., & Murdiyarto, D. (2007). The effect of afforestation on water infiltration in the tropics: A systematic review and meta-analysis. *Forest Ecology and Management*, *251*, 45–51.
- Ingram, V., Haverhals, M., Petersen, S., Elias, M., Basnett, B., & Sola, P. 2016. Gender and forest, tree and agroforestry value chains—Evidence from literature.
- Intergovernmental Panel on Climate Change. 2015. Climate change 2014: Mitigation of climate change.
- IPES-Food. (2016). From uniformity to diversity: A paradigm shift from industrial agriculture to diversified agroecological systems. International Panel of Experts on Sustainable Food Systems.
- Jack, B. K., & Cardona Santos, E. (2017). The leakage and livelihood impacts of PES contracts: A targeting experiment in Malawi. *Land Use Policy*, *63*, 645–658.
- Jayachandran, S., De Laat, J., Lambin, E. F., Stanton, C. Y., Audy, R., & Thomas, N. E. (2017). Cash for carbon: A randomized trial of payments for ecosystem services to reduce deforestation. *Science*, *357*, 267–273.
- Jezeer, R. E., Verweij, P. A., Santos, M. J., & Boot, R. G. A. (2017). Shaded Coffee and Cocoa—Double Dividend for Biodiversity and Small-scale Farmers. *Ecological Economics*, *140*, 136–145.
- Jose, S. (2009). Agroforestry for ecosystem services and environmental benefits: An overview. *Agroforestry Systems*, *76*, 1–10.
- Kalaba, K. F., Chirwa, P., Syampungani, S., & Ajayi, C. O. (2010). Contribution of agroforestry to biodiversity and livelihoods improvement in rural communities of southern African regions. In T. Tschardtke, C. Leuschner, E. Veldkamp, H. Faust, E. Guhardja & A. Bidin (Eds.), *Tropical rainforests and agroforests under global change: Ecological and socio-economic valuations* (pp. 461–476). Springer Berlin Heidelberg.
- Kim, D.-G., Kirschbaum, M. U. F., & Beedy, T. L. (2016). Carbon sequestration and net emissions of CH₄ and N₂O under agroforestry: Synthesizing available data and suggestions for future studies. *Agriculture, Ecosystems & Environment*, *226*, 65–78.
- Kuntashula, E., & Mungatana, E. (2013). Estimating the causal effect of improved fallows on farmer welfare using robust identification strategies in Chongwe, Zambia. *Agroforestry Systems*, *87*, 1229–1246.
- Lawry, S., Samii, C., Hall, R., Leopold, A., Hornby, D., & Mtero, F. (2014). The impact of land property rights interventions on investment and agricultural productivity in developing countries: A systematic review. *Campbell Systematic Reviews*, *10*, 1–104.
- Leisher, C., Tamsah, G., Booker, F., Day, M., Samberg, L., Prosnitz, D., Agarwal, B., Matthews, E., Roe, D., Russell, D., Sunderland, T., Wilkie, D. (2016). Does the gender composition of forest and fishery management groups affect resource governance and conservation outcomes? A systematic map. *Environmental Evidence*, *5*(1), <http://doi.org/10.1186/s13750-016-0057-8>
- Loevinsohn, M., Sumberg, J., Diagne, A., & Whitfield, S. (2013). Under what circumstances and conditions does adoption of technology result in increased agricultural productivity? A systematic review.
- Lopez-Avila, D., Husain, S., Bhatia, R., Nath, M., & Vinayagam, R. (2017). *Agricultural innovation: an evidence gap map* (3ie Evidence Gap Map Report 12). New Delhi: International Initiative for Impact Evaluation (3ie).
- Mafongoya, P. L., Sileshi, G., & Kuntashula, E. (2006). Managing soil fertility and nutrient cycles through fertilizer trees in southern Africa. In N. Uphoff, A. S. Ball, E. Fernandes, H. Herren, O. Husson, M. Laing,

- C. Palm, J. Pretty, P. Sanchez, N. Sangina & J. Thies (Eds.), *Biological approaches to sustainable soil systems* (pp. 273–289). CRC Press.
- Martin, D. A., Osen, K., Grass, I., Hölscher, D., Tschardtke, T., Wurz, A., & Kreft, H. (2020). Land-use history determines ecosystem services and conservation value in tropical agroforestry. *Conservation Letters*, e12740.
- Masset, E., Haddad, L., Cornelius, A., & Isaza-Castro, J. (2012). Effectiveness of agricultural interventions that aim to improve nutritional status of children: systematic review. *BMJ*, 344, d8222.
- Maxwell, S. L., Fuller, R. A., Brooks, T. M., & Watson, J. E. (2016). Biodiversity: The ravages of guns, nets and bulldozers. *Nature*, 536, 143–145.
- Mbow, C., Smith, P., Skole, D., Duguma, L., & Bustamante, M. (2014a). Achieving mitigation and adaptation to climate change through sustainable agroforestry practices in Africa. *Current Opinion in Environmental Sustainability*, 6, 8–14.
- Mbow, C., Van Noordwijk, M., Luedeling, E., Neufeldt, H., Minang, P. A., & Kowero, G. (2014b). Agroforestry solutions to address food security and climate change challenges in Africa. *Current Opinion in Environmental Sustainability*, 6, 61–67.
- Mercer, D. E. (2004). Adoption of agroforestry innovations in the tropics: A review. *Agroforestry Systems*, 61, 311–328.
- Mercer, D. E., & Pattanayak, S. K. (2003). Agroforestry adoption by smallholders. In E. O. Sills & K. L. Abt (Eds.), *Forests in a market economy* (pp. 283–299). Springer Netherlands.
- Miller, D. C., Muñoz-Mora, J. C., & Christiaensen, L. (2017). Prevalence, economic contribution, and determinants of trees on farms across Sub-Saharan Africa. *Forest Policy and Economics*, 84, 47–61.
- Miller, D. C., Ordonez, P. J., Baylis, K., Hughes, K., & Rana, P. (2017). Protocol for an evidence and gap map The impacts of agroforestry on agricultural productivity, ecosystem services, and human well-being in low- and middle-income countries: an evidence and gap map. *Campbell Systematic Reviews*, 13, 1–27.
- Miller, D. C., Ordoñez, P. J., Brown, S. E., Forrest, S., Nava, N. J., Hughes, K., & Baylis, K. (2020). The impacts of agroforestry on agricultural productivity, ecosystem services, and human well-being in low-and middle-income countries: An evidence and gap map. *Campbell Systematic Reviews*, 16, e1066.
- Nair, P. K. R. (1985). Classification of agroforestry systems. *Agroforestry Systems*, 3, 97–128.
- Nair, P. R. (1993). *An introduction to agroforestry*. Springer Science & Business Media.
- Nations, U. (2015). Transforming our world: The 2030 agenda for sustainable development. *General Assembly 70 session*.
- Nguyen, H., Herbohn, J., Clendenning, J., Lamb, D., Dressler, W., Vanclay, J., & Firn, J. (2015). What is the available evidence concerning relative performance of different designs of mixed-species plantings for smallholder and community forestry in the tropics? A systematic map protocol. *Environmental Evidence*, 4(15). <http://doi.org/10.1186/s13750-015-0041-8>
- Norgrove, L., & Beck, J. (2016). Biodiversity Function and Resilience in Tropical Agroforestry Systems Including Shifting Cultivation. *Current Forestry Reports*, 2, 62–80.
- O'Neill, J., Tabish, H., Welch, V., Petticrew, M., Pottie, K., Clarke, M., Evans, T., Pardo Pardo, J., Waters, E., White, H., & Tugwell, P. (2014). Applying an equity lens to interventions: Using PROGRESS ensures consideration of socially stratifying factors to illuminate inequities in health. *Journal of Clinical Epidemiology*, 67, 56–64.
- Oya, C., Schaefer, F., Skolidou, D., McCosker, C., & Langer, L. (2017). Effects of certification schemes for agricultural production on socio-economic outcomes in low- and middle-income countries: A systematic review. *Campbell Systematic Reviews*, 2017.
- Pagiola, S., Honey-Rosés, J., & Freire-González, J. (2016). Evaluation of the permanence of land use change induced by payments for environmental services in Quindío, Colombia. *PLOS One*, 11, e0147829.
- Pandey, D. N. (2007). Multifunctional agroforestry systems in India. *Current Science*, 92, 455–463.
- Pandey, V. L., Mahendra Dev, S., & Jayachandran, U. (2016). Impact of agricultural interventions on the nutritional status in South Asia: A review. *Food Policy*, 62, 28–40.
- Pattanayak, S. K., Mercer, E. D., Sills, E., & Yang, J. (2003). Taking stock of agroforestry adoption studies. *Agroforestry Systems*, 53, 173–186.
- Pender, J., Suyanto, S., Kerr, J., & Edward, K. (2008). *Impacts of the Hutan Kamasyarakatan Social Forestry Program in the Sumberjaya watershed, West Lampung District of Sumatra, Indonesia*. IFPRI Discussion Paper. International Food Policy Research Institute.
- Place, F., Adato, M., Hebinck, P., & Omosa, M. (2005). *The impact of agroforestry-based soil fertility replenishment practices on the poor in western Kenya*, International Food Policy Research Institute.
- Pretty, J. (2018). Intensification for redesigned and sustainable agricultural systems. *Science*, 362, 362.
- Pretty, J., & Bharucha, Z. P. (2014). Sustainable intensification in agricultural systems. *Annals of Botany*, 114, 1571–1596.
- Pumariño, L., Sileshi, G. W., Gripenberg, S., Kaartinen, R., Barrios, E., Muchane, M. N., Midega, C., & Jonsso, M. (2015). Effects of agroforestry on pest, disease and weed control: A meta-analysis. *Basic and Applied Ecology*, 16, 573–582.
- Reed, J., van Vianen, J., Foli, S., Clendenning, J., Yang, K., MacDonald, M., Petrokofsky, G., Padoch, C., & Sunderland, T. (2017). Trees for life: The ecosystem service contribution of trees to food production and livelihoods in the tropics. *Forest Policy and Economics*, 84, 62–71.
- Republic of Kenya. (2014). *Forest Policy*. Nairobi, Kenya.
- Rivest, D., Paquette, A., Moreno, G., & Messier, C. (2013). A meta-analysis reveals mostly neutral influence of scattered trees on pasture yield along with some contrasted effects depending on functional groups and rainfall conditions. *Agriculture, Ecosystems & Environment*, 165, 74–79.
- Rosenstock, T. S., Dawson, I. K., Aynekulu, E., Chomba, S., Degrande, A., Fornace, K., Jamnadass, R., Kimaro, A., Kindt, R., Lamanna, C., Malesu, M., Mausch, K., McMullin, S., Murage, P., Namoi, N., Njenga, M., Nyoka, I., Paez Valencia, A. M., Sola, P., ... Steward, P. (2019). A planetary health perspective on agroforestry in sub-Saharan Africa. *One Earth*, 1, 330–344.
- Sharma, N., Bohra, B., Pragma, N., Ciannella, R., Dobie, P., & Lehmann, S. (2016). Bioenergy from agroforestry can lead to improved food security, climate change, soil quality, and rural development. *Food and Energy Security*, 5, 165–183.
- Sileshi, G., Akinnifesi, F. K., Ajayi, O. C., & Place, F. (2008). Meta-analysis of maize yield response to woody and herbaceous legumes in sub-Saharan Africa. *Plant and Soil*, 307, 1–19.
- Sinclair, F. L. (1999). A general classification of agroforestry practice. *Agroforestry Systems*, 46, 161–180.
- Snilstveit, B., Stevenson, J., Langer, L., Tannous, N., Ravat, Z., Nduku, P., Polanin, J., Shemilt, I., Evers, J., & Ferraro, P. J. (2019). Incentives for climate mitigation in the land use sector—The effects of payment for environmental services on environmental and socioeconomic outcomes in low- and middle-income countries: A mixed-methods systematic review. *Campbell Systematic Reviews*, 15, e1045.
- Sukhdev, P. (2018). Smarter metrics will help fix our food system. *Nature*, 558, 7–8.
- The Government Office for Science (2011). *Foresight. The Future of Food and Farming*. London.
- Tierney, M. J., Nielson, D. L., Hawkins, D. G., Roberts, J. T., Findley, M. G., Powers, R. M., Parks, B., Wilson, S. E., & Hicks, R. L. (2011). More dollars than sense: Refining our knowledge of development finance using AidData. *World Development*, 39, 1891–1906.

- Torquebiau, E. F. (2000). A renewed perspective on agroforestry concepts and classification. *Comptes Rendus de l'Académie des Sciences-Series III-Sciences de la Vie*, 323, 1009–1017.
- Van Noordwijk, M., Duguma, L. A., Dewi, S., Leimona, B., Catacutan, D. C., Lusiana, B., Öborn, I., Hairiah, K., & Minang, P. A. (2018). SDG synergy between agriculture and forestry in the food, energy, water and income nexus: Reinventing agroforestry? *Current Opinion in Environmental Sustainability*, 34, 33–42.
- Viechtbauer, W. (2010). Conducting meta-analyses in R with the metafor package. *Journal of Statistical Software*, 36, 1–48.
- Vira, B., Wildburger, C. & Mansourian, S. (Eds.). (2015). *Forests and food: Addressing hunger and nutrition across sustainable landscapes*. Open Book Publishers.
- Waldron, A., Garrity, D., Malhi, Y., Girardin, C., Miller, D. C., & Seddon, N. (2017). Agroforestry can enhance food security while meeting other sustainable development goals. *Tropical Conservation Science*, 10, 1–6.
- Waldron, A., Justicia, R., & Smith, L. E. (2015). Making biodiversity-friendly cocoa pay: combining yield, certification, and REDD for shade management. *Ecological Applications*, 25, 361–372.
- Weiser, S. D., Bukusi, E. A., Steinfeld, R. L., Frongillo, E. A., Weke, E., Dworkin, S. L., Pusateri, K., Shiboski, S., Scow, K., & Butler, L. M. (2015). Shamba maisha: Randomized controlled trial of an agricultural and finance intervention to improve HIV health outcomes in Kenya. *AIDS*, 29, 1889–1894.
- World Agroforestry Centre. (2017). Corporate Strategy 2017–2020—Transforming Lives and Landscapes with Trees.
- World Bank. (2016). World Bank Group Climate Change Action Plan 2016–2020. Washington D.C.
- Zomer, R. J., Trabucco, A., Coe, R., Place, F., Van Noordwijk, M., & Xu, J. (2014). *Trees on farms: An update and reanalysis of agroforestry's global extent and socio-ecological characteristics* (World Agroforestry Center Working Paper, 179).

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section.

How to cite this article: Castle, S. E., Miller, D. C., Ordóñez, P. J., Baylis, K., & Hughes, K. The impacts of agroforestry interventions on agricultural productivity, ecosystem services, and human well-being in low- and middle-income countries: A systematic review. *Campbell Systematic Reviews*. 2021;e1167. <https://doi.org/10.1002/cl2.1167>

APPENDIX A: DETAILED SEARCH STRATEGY

Electronic searches

The databases that we searched for publications were:

- SCOPUS
- EBSCO: Econlit

- OVID: Agricola
- Web of Science: Core Collection
- Web of Science: CAB Abstracts and Global Health
- AGRIS

The search terms used in each database are shown in Table A1 (constructed using the terms from CAB thesaurus and using the research group's specific knowledge). Each search string included each of the agroforestry practices from Table 3. These terms and search strings were modified through a scoping exercise in Web of Science, SCOPUS, and EBSCO, where the search terms were used, and the results were evaluated by analyzing the relevance of the first 50 studies. We note that the intervention types are more generic, including topics well beyond agroforestry, so our search focused on practices. We included an LMIC filter to identify only studies in relevant countries, shown in the last row of Table 5, which was used in Lopez-Avila et al. (2017) based on World Bank country classification by income in 2017, when our search was conducted. Specific search strings for the databases are shown in Table A2. In conducting this SR, we identified errors in the search strings used in the earlier EGM for the Scopus and Agricola databases. We corrected the search strings for these databases in this SR, and we researched and screened the additional studies that we had missed. This process allowed us to identify three additional impact evaluations in the Agricola database (no additional relevant studies were identified in the Scopus database).

Searching other resources

Additionally, in order to identify the existing grey literature, the websites of various organizations that are likely to produce published and unpublished research were searched. The list of relevant research organizations (Table A3) was constructed from cross-validation of websites listed in the systematic mapping protocols of agroforestry related studies (e.g., Bottrill et al., 2014; Leisher et al. 2016; Nguyen et al. 2015). This list was validated with the external EGM advisory group. To optimize the scope of the search while ensuring transparency in our methods, we followed the approach developed by Haddaway et al. (2017), which allowed us to search multiple websites simultaneously and to extract the relevant information from each website into a single database. Finally, we also contacted key informants within 3ie, ICRAF, and other relevant organizations for identification of additional relevant literature for screening and inclusion. For the organizational websites, we used a simplified search string with our primary keyword “agroforestry.”

TABLE A1 Search terms by intervention and outcomes

Category	Terms
Practices	("agroforest*" OR "agriforest*" OR "agro-forest*" OR agrosilvicultur* OR agrisilvicultur* OR "improved fallow*" OR "shade tree*" OR "rotational tree fallow*" OR parkland* OR "multipurpose tree*" OR "tree garden*" OR "forest garden" OR "alley cropping" OR intercropping OR "shifting cultivation" OR shelterbelt* OR "natural vegetation strip*" OR "wind break*" OR "sloping agricultural land technology" OR "hedgerows" OR "hedge cropping" OR silvopastoral* OR silvipastoral* OR "fodder tree*" OR "living fence*" OR "integrated animal and wood production" OR "trees on pasture" OR agrosilvopastoral* OR "integrated production of animals, crops and wood" OR "tree-crop-livestock" OR "apiculture with trees" OR entomoforestry OR "aqua-silvo-fisher*" OR "multi-purpose tree lot*" OR "tree* on farms" OR "on-farm tree*" OR "woody hedgerows" OR "wooded pastures produce" OR "fertili*er trees" OR "shade species" OR "shade-grown" OR "alternative agriculture" OR "tree-based system*" OR "tree fallow*" OR "planted fallow*" OR woodlot* OR "boundary planting" OR "mixed trees and crops" OR "conservation agriculture with trees" OR "farmer managed natural regeneration" OR homegarden OR "fodder shrub*" OR "multi-strata systems" OR "nitrogen fixing trees" OR "vegetative field strips")
Study designs	AND("impact" OR "outcome" OR "result" OR "effect*" OR "intervention" OR "evaluation" OR "assessment" OR "*effectiveness" OR "cost-benefit" OR "cost benefit" OR "efficacy" OR "systematic review" OR "field trial" OR "observational stud*" OR "trial" OR "random* control* trial*" OR "random* trial*" OR RCT OR "propensity score matching" OR PSM OR "regression discontinuity design" OR RDD OR "difference in difference*" OR matching OR (random* adj3 allocat*) OR "instrumental variable*" OR IV OR evaluation OR assessment OR "comparison group" OR counterfactual OR "counter factual" OR counter-factual OR quasi-experimental OR quasiexperimental OR ((quantitative or experiment*) adj3 (design OR study OR analysis)))
LMIC filter	(Afghanistan OR Angola OR Albania OR "American Samoa" OR Armenia OR Armenian OR Azerbaijan OR Bangladesh OR Belarus OR Belize OR Benin OR Bolivia OR Bosnia OR Herzegovina OR Botswana OR Brazil OR Bulgaria OR "Burkina Faso" OR "Burkina Faso" OR Burundi OR Urundi OR "Cabo Verde" OR "Cape Verde" OR Cambodia OR Cameroon OR Camerons OR Cameron OR Camerons OR "Central African Republic" OR Chad OR China OR Colombia OR Comoros OR "Comoro Islands" OR Comores OR Congo OR "Costa Rica" OR Cuba OR Zaire OR "Cote d'Ivoire" OR "Ivory Coast" OR Djibouti OR Dominica* OR "East Timor" OR "East Timor" OR "Timor Leste" OR Ecuador OR Egypt OR "United Arab Republic" OR "El Salvador" OR Eritrea OR Ethiopia OR Fiji OR Gabon OR Gambia OR Gaza OR "Georgia Republic" OR "Georgian Republic" OR Ghana OR Grenada OR Guatemala OR Guinea OR Guiana OR Guyana OR Haiti OR Honduras OR India OR Indonesia OR Iran OR Iraq OR Kazakhstan OR Kenya OR Kiribati OR Korea OR Kosovo OR Kyrgyzstan OR Kirghizia OR "Kyrgyz Republic" OR Kirghiz OR Kirgizstan OR "Lao PDR" OR Laos OR Lebanon OR Lesotho OR Liberia OR Libya OR Macedonia OR Madagascar OR "Malagasy Republic" OR Malawi OR Malaysia OR Maldives OR "Marshall Islands" OR Mali OR Mauritania OR Mauritius OR "Agalega Islands" OR Mexico OR Micronesia OR Moldova OR Moldovia OR Moldovian OR Mongolia OR Montenegro OR Morocco OR Ifni OR Mozambique OR Myanmar OR Myanma OR Burma OR Namibia OR Nepal OR Nicaragua OR Niger OR Nigeria OR Pakistan OR Palau OR Palestine OR Panama OR Paraguay OR Peru OR Philippines OR Philipines OR Phillipines OR Phillipines OR Romania OR Rwanda OR Ruanda OR Samoa OR "Samoan Islands" OR "Sao Tome" OR Senegal OR Serbia OR "Sierra Leone" OR "Sri Lanka" OR "Solomon Islands" OR Somalia OR "South Africa" OR "St Lucia" OR "St Vincent" OR Grenadines OR Sudan OR Suriname OR Swaziland OR Syria OR Tajikistan OR Tadjhikistan OR Tadjikistan OR Tadjhik OR Tanzania OR Thailand OR Tonga OR Togo OR "Togolese Republic" OR Tunisia OR Turkey OR Turkmenistan OR Tuvalu OR Uganda OR Ukraine OR Uzbekistan OR Uzbek OR Vanuatu OR "New Hebrides" OR Vietnam OR "Viet Nam" OR "West Bank" OR Yemen OR Zambia OR Zimbabwe)

TABLE A2 Database specific search strings

Electronic database	Search string	Dates searched	Number sources returned
Web of Science—Core Collection and CAB Abstracts and Global Health	<p>(TS = (("agroforest*" OR "agriforest*" OR "agro-forest*" OR agrosilvicultur* OR agrisilvicultur* OR "improved fallow*" OR "shade tree*" OR "rotational tree fallow*" OR parkland* OR "multipurpose tree*" OR "tree garden*" OR "forest garden" OR "alley cropping" OR (intercropping AND ("tree*" OR "perennial*"))) OR "shifting cultivation" OR shelterbelt* OR "natural vegetation strip*" OR "wind break*" OR "sloping agricultural land technology" OR hedgerows OR "hedge cropping" OR silvopastoral* OR silvipastoral* OR "fodder tree*" OR "living fence*" OR "integrated animal and wood production" OR "trees on pasture" OR agrosilvopastoral* OR "integrated production of animals, crops and wood" OR "tree-crop-livestock" OR "apiculture with trees" OR entomoforestry OR "aqua-silvo-fisher*" OR "multi-purpose tree lot*" OR "tree* on farms" OR "on-farm tree*" OR "woody hedgerows" OR "wooded pastures produce" OR "fertili*er trees" OR "shade species" OR "shade-grown" OR "alternative agriculture" OR "tree-based system*" OR "tree fallow*" OR "planted fallow*" OR woodlot* OR "boundary planting" OR "mixed trees and crops" OR "conservation agriculture with trees" OR "farmer managed natural regeneration" OR homegarden OR "fodder shrub*" OR "multi-strata systems" OR "nitrogen fixing trees" OR "vegetative field strips")</p> <p>AND</p> <p>(impact OR outcome OR result OR effect* OR intervention OR evaluation OR assessment OR *effectiveness OR cost-benefit OR cost benefit OR efficacy OR systematic review OR field trial OR observational stud* OR trial OR "random* control* trial*" OR "random*trial*" OR RCT OR "propensity score matching" OR PSM OR "regression discontinuity design" OR RDD OR "difference in difference*" OR matching OR "instrumental variable*" OR IV OR evaluation OR assessment OR "comparison group" OR counterfactual OR "counter factual" OR counterfactual OR quasi-experimental OR quasiexperimental)</p> <p>AND</p> <p>(Afghanistan OR Angola OR Albania OR "American Samoa" OR Armenia OR Armenian OR Azerbaijan OR Bangladesh OR Belarus OR Belize OR Benin OR Bolivia OR Bosnia OR Herzegovina OR Botswana OR Brazil OR Bulgaria OR "Burkina Faso" OR "Burkina Fasso" OR Burundi OR Urundi OR "Cabo Verde" OR "Cape Verde" OR Cambodia OR Cameroon OR Cameroons OR Cameron OR Camerons OR "Central African Republic" OR Chad OR China OR Colombia OR Comoros OR "Comoro Islands" OR Comores OR Congo OR "Costa Rica" OR Cuba OR Zaire OR "Cote d'Ivoire" OR "Ivory Coast" OR Djibouti OR Dominica* OR "East Timor" OR "East Timur" OR "Timor Leste" OR Ecuador OR Egypt OR "United Arab Republic" OR "El Salvador" OR Eritrea OR Ethiopia OR Fiji OR Gabon OR Gambia OR Gaza OR "Georgia Republic" OR "Georgian Republic" OR Ghana OR Grenada OR Guatemala OR Guinea OR Guiana OR Guyana OR Haiti OR Honduras OR India OR Indonesia OR Iran OR Iraq OR Kazakhstan OR Kenya OR Kiribati OR Korea OR Kosovo OR Kyrgyzstan OR Kirghizia OR "Kyrgyz Republic" OR Kirghiz OR Kirgizstan OR "Lao PDR" OR Laos OR Lebanon OR Lesotho OR Liberia OR Libya OR Macedonia OR Madagascar OR "Malagasy Republic" OR Malawi OR Malaysia OR Maldives OR "Marshall Islands" OR Mali OR</p>	January 1, 2000 through October 20, 2017	Core Collection = 3,877 CAB Abstracts and Global Health = 12,090

TABLE A2 (Continued)

Electronic database	Search string	Dates searched	Number sources returned
	<p>Mauritania OR Mauritius OR "Agalega Islands" OR Mexico OR Micronesia OR Moldova OR Moldavia OR Moldovan OR Mongolia OR Montenegro OR Morocco OR Ifni OR Mozambique OR Myanmar OR Myanma OR Burma OR Namibia OR Nepal OR Nicaragua OR Niger OR Nigeria OR Pakistan OR Palau OR Palestine OR Panama OR Paraguay OR Peru OR Philippines OR Philipines OR Phillipines OR Phillippines OR Romania OR Rwanda OR Ruanda OR Samoa OR "Samoan Islands" OR "Sao Tome" OR Senegal OR Serbia OR "Sierra Leone" OR "Sri Lanka" OR "Solomon Islands" OR Somalia OR "South Africa" OR "St Lucia" OR "St Vincent" OR Grenadines OR Sudan OR Suriname OR Swaziland OR Syria OR Tajikistan OR Tadjhikistan OR Tadjikistan OR Tadjhik OR Tanzania OR Thailand OR Tonga OR Togo OR "Togolese Republic" OR Tunisia OR Turkey OR Turkmenistan OR Tuvalu OR Uganda OR Ukraine OR Uzbekistan OR Uzbek OR Vanuatu OR "New Hebrides" OR Vietnam OR "Viet Nam" OR "West Bank" OR Yemen OR Zambia OR Zimbabwe)))</p> <p>AND</p> <p>LANGUAGE: (English)</p> <p>Indexes=SCI-EXPANDED, SSCI, A&HCI, CPCI-S, CPCI-SSH, BKCI-S, BKCI-SSH, ESCI, CCR-EXPANDED, IC</p> <p>Timespan=2000-2017</p>		
Scopus	<p>TITLE-ABS-KEY(("agroforest*" OR "agriforest*" OR "agroforest*" OR "agrosilvicultur*" OR "agrisilvicultur*" OR "improved fallow*" OR "shade tree*" OR "rotational tree fallow*" OR "parkland*" OR "multipurpose tree*" OR "tree garden*" OR "forest garden" OR "alley cropping" OR (intercropping AND ("tree*" OR "perennial*")) OR "shifting cultivation" OR shelterbelt* OR "natural vegetation strip*" OR "wind break*" OR "sloping agricultural land technology" OR "hedgerows" OR "hedge cropping" OR silvopastoral* OR silvipastoral* OR "fodder tree*" OR "living fence*" OR "integrated animal and wood production" OR "trees on pasture" OR agrosilvopastoral* OR "integrated production of animals, crops and wood" OR "tree-crop-livestock" OR "apiculture with trees" OR entomoforestry OR "aqua-silvo-fisher*" OR "multi-purpose tree lot*" OR "tree* on farms" OR "on-farm tree*" OR "woody hedgerows" OR "wooded pastures produce" OR "fertili*er trees" OR "shade species" OR "shade-grown" OR "alternative agriculture" OR "tree-based system*" OR "tree fallow*" OR "planted fallow*" OR woodlot* OR "boundary planting" OR "mixed trees and crops" OR "conservation agriculture with trees" OR "farmer managed natural regeneration" OR homegarden OR "fodder shrub*" OR "multi-strata systems" OR "nitrogen fixing trees" OR "vegetative field strips") AND (impact OR outcome OR result OR effect* OR intervention OR evaluation OR assessment OR *effectiveness OR cost-benefit OR (cost AND benefit) OR efficacy OR (systematic AND review) OR (field AND trial) OR (observational AND stud*) OR trial OR "random* control* trial*" OR "random* trial*" OR rct OR "propensity score matching" OR psm OR "regression discontinuity design" OR rdd OR "difference in difference*" OR matching OR ((random* W/3 allocat*)) OR "instrumental variable*" OR iv OR evaluation OR assessment OR "comparison group" OR counterfactual OR "counterfactual" OR counter-factual OR quasi-experimental OR</p>	<p>January 1, 2000 through October 20, 2017</p>	<p>45* (*Corrected search yielded 5,254 results. 2,685 after removing duplicates. 2,555 excluded on title/abstract. 130 excluded on full text. Zero additional included papers were identified.)</p>

TABLE A2 (Continued)

Electronic database	Search string	Dates searched	Number sources returned
	<p>quasiexperimental OR ((quantitative OR experiment*) W/ 3 (design OR study OR analysis))) AND (Afghanistan OR Angola OR Albania OR "American Samoa" OR Armenia OR Armenian OR Azerbaijan OR Bangladesh OR Belarus OR Belize OR Benin OR Bolivia OR Bosnia OR Herzegovina OR Botswana OR Brazil OR Bulgaria OR "Burkina Faso" OR "Burkina Fasso" OR Burundi OR Urundi OR "Cabo Verde" OR "Cape Verde" OR Cambodia OR Cameroon OR Cameroons OR Cameron OR Camerons OR "Central African Republic" OR Chad OR China OR Colombia OR Comoros OR "Comoro Islands" OR Comores OR Congo OR "Costa Rica" OR Cuba OR Zaire OR "Cote d'Ivoire" OR "Ivory Coast" OR Djibouti OR Dominica* OR "East Timor" OR "East Timur" OR "Timor Leste" OR Ecuador OR Egypt OR "United Arab Republic" OR "El Salvador" OR Eritrea OR Ethiopia OR Fiji OR Gabon OR Gambia OR Gaza OR "Georgia Republic" OR "Georgian Republic" OR Ghana OR Grenada OR Guatemala OR Guinea OR Guiana OR Guyana OR Haiti OR Honduras OR India OR Indonesia OR Iran OR Iraq OR Kazakhstan OR Kenya OR Kiribati OR Korea OR Kosovo OR Kyrgyzstan OR Kirghizia OR "Kyrgyz Republic" OR Kirghiz OR Kirgizstan OR "Lao PDR" OR Laos OR Lebanon OR Lesotho OR Liberia OR Libya OR Macedonia OR Madagascar OR "Malagasy Republic" OR Malawi OR Malaysia OR Maldives OR "Marshall Islands" OR Mali OR Mauritania OR Mauritius OR "Agalega Islands" OR Mexico OR Micronesia OR Moldova OR Moldovia OR Moldovan OR Mongolia OR Montenegro OR Morocco OR Ifni OR Mozambique OR Myanmar OR Myanma OR Burma OR Namibia OR Nepal OR Nicaragua OR Niger OR Nigeria OR Pakistan OR Palau OR Palestine OR Panama OR Paraguay OR Peru OR Philippines OR Philipines OR Phillipines OR Phillipines OR Romania OR Rwanda OR Ruanda OR Samoa OR "Samoan Islands" OR "Sao Tome" OR Senegal OR Serbia OR "Sierra Leone" OR "Sri Lanka" OR "Solomon Islands" OR Somalia OR "South Africa" OR "St Lucia" OR "St Vincent" OR Grenadines OR Sudan OR Suriname OR Swaziland OR Syria OR Tajikistan OR Tadjhikistan OR Tadjikistan OR Tadjhik OR Tanzania OR Thailand OR Tonga OR Togo OR "Togolese Republic" OR Tunisia OR Turkey OR Turkmenistan OR Tuvalu OR Uganda OR Ukraine OR Uzbekistan OR Uzbek OR Vanuatu OR "New Hebrides" OR Vietnam OR "Viet Nam" OR "West Bank" OR Yemen OR Zambia OR Zimbabwe)) AND (PUBYEAR > 1999) AND (LANGUAGE(English))</p>		
EBSCO: Econlit	<p>((("agroforest*" OR "agriforest*" OR "agro-forest*" OR agrosilvicultur* OR agrisilvicultur* OR "improved fallow*" OR "shade tree*" OR "rotational tree fallow*" OR parkland* OR "multipurpose tree*" OR "tree garden*" OR "forest garden" OR "alley cropping" OR intercropping OR "shifting cultivation" OR shelterbelt* OR "natural vegetation strip*" OR "wind break*" OR "sloping agricultural land technology" OR hedgerows OR "hedge cropping" OR silvopastoral* OR silvipastoral* OR "fodder tree*" OR "living fence*" OR "integrated animal and wood production" OR "trees on pasture" OR agrosilvopastoral* OR "integrated production of animals, crops and wood" OR "tree-crop-livestock" OR "apiculture with trees" OR entomoforestry OR "aqua-silvo-fisher*" OR "multi-purpose tree lot*" OR "tree* on farms" OR "on-farm tree*"))</p>	January 1, 2000 through October 21, 2017	177

TABLE A2 (Continued)

Electronic database	Search string	Dates searched	Number sources returned
	<p>OR "woody hedgerows" OR "wooded pastures produce" OR "fertili*er trees" OR "shade species" OR "shade-grown" OR "alternative agriculture" OR "tree-based system*" OR "tree fallow*" OR "planted fallow*" OR woodlot* OR "boundary planting" OR "mixed trees and crops" OR "conservation agriculture with trees" OR "farmer managed natural regeneration" OR homegarden OR "fodder shrub*" OR "multi-strata systems" OR "nitrogen fixing trees" OR "vegetative field strips") AND (impact OR outcome OR result OR effect* OR intervention OR evaluation OR assessment OR *effectiveness OR cost-benefit OR cost benefit OR efficacy OR systematic review OR field trial OR observational stud* OR trial OR "random* control* trial*" OR "random* trial*" OR RCT OR "propensity score matching" OR PSM OR "regression discontinuity design" OR RDD OR "difference in difference*" OR matching OR ((random* W3 allocat*)) OR "instrumental variable*" OR IV OR evaluation OR assessment OR "comparison group" OR counterfactual OR "counter factual" OR counterfactual OR quasi-experimental OR quasiexperimental OR ((quantitative or experiment*) W3 (design OR study OR analysis))) AND (Afghanistan OR Angola OR Albania OR "American Samoa" OR Armenia OR Armenian OR Azerbaijan OR Bangladesh OR Belarus OR Belize OR Benin OR Bolivia OR Bosnia OR Herzegovina OR Botswana OR Brazil OR Bulgaria OR "Burkina Faso" OR "Burkina Fasso" OR Burundi OR Urundi OR "Cabo Verde" OR "Cape Verde" OR Cambodia OR Cameroon OR Camerons OR Cameron OR Camerons OR "Central African Republic" OR Chad OR China OR Colombia OR Comoros OR "Comoro Islands" OR Comores OR Congo OR "Costa Rica" OR Cuba OR Zaire OR "Cote d'Ivoire" OR "Ivory Coast" OR Djibouti OR Dominica* OR "East Timor" OR "East Timur" OR "Timor Leste" OR Ecuador OR Egypt OR "United Arab Republic" OR "El Salvador" OR Eritrea OR Ethiopia OR Fiji OR Gabon OR Gambia OR Gaza OR "Georgia Republic" OR "Georgian Republic" OR Ghana OR Grenada OR Guatemala OR Guinea OR Guiana OR Guyana OR Haiti OR Honduras OR India OR Indonesia OR Iran OR Iraq OR Kazakhstan OR Kenya OR Kiribati OR Korea OR Kosovo OR Kyrgyzstan OR Kirghizia OR "Kyrgyz Republic" OR Kirghiz OR Kirgizstan OR "Lao PDR" OR Laos OR Lebanon OR Lesotho OR Liberia OR Libya OR Macedonia OR Madagascar OR "Malagasy Republic" OR Malawi OR Malaysia OR Maldives OR "Marshall Islands" OR Mali OR Mauritania OR Mauritius OR "Agalega Islands" OR Mexico OR Micronesia OR Moldova OR Moldavia OR Moldovan OR Mongolia OR Montenegro OR Morocco OR Ifni OR Mozambique OR Myanmar OR Myanma OR Burma OR Namibia OR Nepal OR Nicaragua OR Niger OR Nigeria OR Pakistan OR Palau OR Palestine OR Panama OR Paraguay OR Peru OR Philippines OR Philipines OR Phillipines OR Phillippines OR Romania OR Rwanda OR Ruanda OR Samoa OR "Samoan Islands" OR "Sao Tome" OR Senegal OR Serbia OR "Sierra Leone" OR "Sri Lanka" OR "Solomon Islands" OR Somalia OR "South Africa" OR "St Lucia" OR "St Vincent" OR Grenadines OR Sudan OR Suriname OR Swaziland OR Syria OR Tajikistan OR Tadjhikistan OR Tadjikistan OR Tadjhik OR Tanzania OR Thailand OR Tonga OR Togo OR "Togolese Republic" OR Tunisia OR</p>		

TABLE A2 (Continued)

Electronic database	Search string	Dates searched	Number sources returned
	Turkey OR Turkmenistan OR Tuvalu OR Uganda OR Ukraine OR Uzbekistan OR Uzbek OR Vanuatu OR "New Hebrides" OR Vietnam OR "Viet Nam" OR "West Bank" OR Yemen OR Zambia OR Zimbabwe)		
OVID: Agricola	<p>("agroforest*" or "agriforest*" or "agro-forest*" or agrosilvicultur* or agrisilvicultur* or "improved fallow*" or "shade tree*" or "rotational tree fallow*" or parkland* or "multipurpose tree*" or "tree garden*" or "forest garden" or "alley cropping" or (intercropping AND ("tree*" OR "perennial*")) or "shifting cultivation" or shelterbelt* or "natural vegetation strip*" or "wind break*" or "sloping agricultural land technology" or hedgerows or "hedge cropping" or silvopastoral* or silvipastoral* or "fodder tree*" or "living fence*" or "integrated animal and wood production" or "trees on pasture" or agrosilvopastoral* or "integrated production of animals, crops and wood" or "tree-crop-livestock" or "apiculture with trees" or entomoforestry or "aqua-silvo-fisher*" or "multi-purpose tree lot*" or "tree* on farms" or "on-farm tree*" or "woody hedgerows" or "wooded pastures produce" or "fertili*er trees" or "shade species" or "shade-grown" or "alternative agriculture" or "tree-based system*" or "tree fallow*" or "planted fallow*" or woodlot* or "boundary planting" or "mixed trees and crops" or "conservation agriculture with trees" or "farmer managed natural regeneration" or homegarden or "fodder shrub*" or "multi-strata systems" or "nitrogen fixing trees")</p> <p>limit to (english language and yr = "2000 - 2017")</p> <p>AND</p> <p>(impact OR outcome OR result OR effect* OR intervention OR evaluation OR assessment OR *effectiveness OR cost-benefit OR cost benefit OR efficacy OR systematic review OR field trial OR observational stud* OR trial OR "random* control* trial*" OR "random* trial*" OR RCT OR "propensity score matching" OR PSM OR "regression discontinuity design" OR RDD OR "difference in difference*" OR matching OR ((random* AND allocat*)) OR "instrumental variable*" OR IV OR evaluation OR assessment OR "comparison group" OR counterfactual OR "counter factual" OR counter-factual OR quasi-experimental OR quasiexperimental OR ((quantitative or experiment*) AND (design OR study OR analysis)))</p> <p>limit to (english language and yr = "2000 - 2017")</p> <p>AND</p> <p>(Afghanistan or Angola or Albania or "American Samoa" or Armenia or Armenian or Azerbaijan or Bangladesh or Belarus or Belize or Benin or Bolivia or Bosnia or Herzegovina or Botswana or Brazil or Bulgaria or "Burkina Faso" or "Burkina Fasso" or Burundi or Urundi or "Cabo Verde" or "Cape Verde" or Cambodia or Cameroon or Cameroons or Cameron or Camerons or "Central African Republic" or Chad or China or Colombia or Comoros or "Comoro Islands" or Comores or Congo or "Costa Rica" or Cuba or Zaire or "Cote d'Ivoire" or "Ivory Coast" or Djibouti or Dominica* or "East Timor" or "East Timur" or "Timor Leste" or Ecuador or Egypt or "United Arab Republic" or "El Salvador" or Eritrea or Ethiopia or Fiji or Gabon or Gambia or Gaza or "Georgia Republic" or "Georgian Republic" or Ghana or Grenada or Guatemala or Guinea or Guiana or Guyana or Haiti or Honduras or India or Indonesia or Iran or Iraq or Kazakhstan or Kenya</p>	January 1, 2000 through October 20, 2017	63* (*Corrected search yielded 1,620 results. 352 after removing duplicates. 322 excluded on title/abstract. 27 excluded on full text. 3 additional included papers were identified.)

TABLE A2 (Continued)

Electronic database	Search string	Dates searched	Number sources returned
	<p>or Kiribati or Korea or Kosovo or Kyrgyzstan or Kirghizia or "Kyrgyz Republic" or Kirghiz or Kirgizstan or "Lao PDR" or Laos or Lebanon or Lesotho or Liberia or Libya or Macedonia or Madagascar or "Malagasy Republic" or Malawi or Malaysia or Maldives or "Marshall Islands" or Mali or Mauritania or Mauritius or "Agalega Islands" or Mexico or Micronesia or Moldova or Moldovia or Moldovian or Mongolia or Montenegro or Morocco or Ifni or Mozambique or Myanmar or Myanma or Burma or Namibia or Nepal or Nicaragua or Niger or Nigeria or Pakistan or Palau or Palestine or Panama or Paraguay or Peru or Philippines or Philipines or Phillipines or Phillipines or Romania or Rwanda or Ruanda or Samoa or "Samoan Islands" or "Sao Tome" or Senegal or Serbia or "Sierra Leone" or "Sri Lanka" or "Solomon Islands" or Somalia or "South Africa" or "St Lucia" or "St Vincent" or Grenadines or Sudan or Suriname or Swaziland or Syria or Tajikistan or TadzhiKistan or Tadjikistan or TadzhiK or Tanzania or Thailand or Tonga or Togo or "Togolese Republic" or Tunisia or Turkey or Turkmenistan or Tuvalu or Uganda or Ukraine or Uzbekistan or Uzbek or Vanuatu or "New Hebrides" or Vietnam or "Viet Nam" or "West Bank" or Yemen or Zambia or Zimbabwe)</p> <p>limit to (english language and yr = "2000 - 2017")</p>		
AGRIS	<p>(agroforest* OR agriforest* OR agro-forest* OR agrosilvicultur* OR agrisilvicultur* OR "improved fallow" OR "shade tree" OR "rotational tree fallow" OR parkland OR "multipurpose tree" OR "tree garden" OR "forest garden" OR "alley cropping" OR "shifting cultivation" OR shelterbelt OR "natural vegetation strip" OR "wind break" OR "sloping agricultural land technology" OR hedgerow* OR "hedge cropping" OR silvopastoral OR silvipastoral OR "fodder tree" OR "living fence" OR "trees on pasture" OR agrosilvopastoral OR "tree-crop-livestock" OR "apiculture with trees" OR entomoforestry OR "aqua-silvo-fishery" OR "multi-purpose tree lot" OR "trees on farms" OR "on-farm tree" OR "woody hedgerows" OR "wooded pastures" OR "fertilizer trees" OR "shade species" OR "shade-grown" OR "alternative agriculture" OR "tree-based system" OR "tree fallow" OR "planted fallow" OR woodlot* OR "boundary planting" OR "mixed trees and crops" OR "conservation agriculture with trees" OR "farmer managed natural regeneration" OR homegarden* OR "fodder shrub" OR "multi-strata systems" OR "nitrogen fixing trees" OR intercropping) AND (impact OR outcome OR result OR effect OR intervention OR evaluation OR assessment OR effectiveness OR cost-benefit OR efficacy OR "systematic review" OR "field trial" OR "observational study" OR trial OR "random control trial" OR "random trial" OR RCT OR "propensity score matching" OR PSM OR "regression discontinuity design" OR RDD OR "difference in difference" OR matching OR "instrumental variable" OR IV OR evaluation OR assessment OR "comparison group" OR counterfactual OR counter-factual OR quasi-experimental OR quasiexperimental)publication</p> <p>date = 2000-2017, language = English</p>	<p>January 1, 2000 through October 20, 2017</p>	523

TABLE A3 List of websites from relevant organizations

Organization	Website	# of sources returned
Australian Centre for International Agricultural Research (ACIAR)	http://aciarc.gov.au/aboutus	11
Center for International Forestry Research (CIFOR)	http://www.cifor.org	126
Economy and Environment Program for Southeast Asia (EEPSEA)	www.eepsea.org	3
Food and Agriculture Organization (FAO)	http://www.fao.org	963
French Agricultural Research Centre for International Development	www.cirad.fr	99
GFIS	www.GFIS.net	380
IDEAS RePEc (Research Papers in Economics)	https://ideas.repec.org	297
Inter-American Development Bank (IADB)	www.iadb.org	3
International Food Policy Research Institute Library (IFPRI)	http://library.ifpri.info/	36
International Institute for Environment and Development	http://www.iied.org	8
International Impact Initiative (3ie)	http://www.3ieimpact.org/	0
International Tropical Timber Organization	www.itto.int	8
Overseas Development Institute	https://www.odi.org/	208
Tropical Agricultural Research and Higher Education Center (CATIE)	http://www.catie.ac.cr/en/	75
United Nations Development Programme (UNDP)	http://www.undp.org	6
United States Agency for International Development (USAID)	http://www.usaid.gov	51
USAID Development Experience Clearinghouse	dec.usaid.gov	505
World Agroforestry Center (ICRAF)	www.worldagroforestry.org	577
World Resources Institute	http://www.wri.org/	5