ABSTRACT

Title of Dissertation: EFFECTIVELY EVALUATING ENVIRONMENTAL, SOCIAL, AND ECONOMIC OUTCOMES IN VOLUNTARY SUSTAINABILITY PROGRAMS: LESSONS FROM LAOS

Rebecca Traldi, Doctor of Philosophy, 2022

Dissertation directed by: Associate Professor, Dr. Julie A. Silva, Department of Geographical Sciences

Voluntary sustainability programs (VSPs) are a subset of environmental interventions which rely on participants' willingness to engage, rather than mandatory regulation. VSPs have been a central component of sustainable development and environmental mitigation strategies for decades, with significant investments from nongovernmental organizations (NGOs), multilaterals, and the private sector. VSPs typically aim to positively influence environmental, economic, and social outcomes, although program-specific priorities often result in an uneven focus across these three domains (also known as the three pillars of sustainability). Despite their popularity, questions regarding the value of VSPs remain unanswered. Assessments of VSPs typically do not eliminate rival explanations for program outcomes when evaluating their successes and failures, thus limiting our understanding of their effectiveness.

This dissertation addresses this gap by investigating socioeconomic and environmental outcomes for agriculture and forestry VSPs. Mixed methods including systematic review, inverse

probability-of-treatment weighted regression (IPWR), and inequality and polarization decomposition provide insights both at a global level, and for two national case studies in Lao People's Democratic Republic (hereafter Laos). A wide range of datasets inform the analysis, including nationally representative poverty and expenditure surveys and land-use land cover estimates derived from remotely sensed imagery. By exploring a variety of VSPs – including agricultural and forestry voluntary sustainability standards and sustainable development projects – the study acknowledges the context-specific nature of VSP impact, while also drawing generalizable insights relevant for different types of interventions.

The research findings presented in this dissertation elucidate the degree to which VSPs deliver on stated goals and objectives. First, a systematic literature review reveals that the evidence base for VSP impact remains limited, with some geographies, sustainability outcomes, and project types receiving more inquiry and evaluation than others. Second, an IPWR analysis suggests that agriculture and forestry VSPs have achieved some success in generating positive outcomes – specifically, for poverty and forest cover. However, variations in project focus and design bring different results. For example, food security and livelihoods programs which prioritize local socioeconomic well-being can generate significant co-benefits for environmental outcomes, and resource management projects can positively impact forest cover. Conversely, the forest management projects considered here do not achieve significant benefits for poverty or forest cover – presumably due to challenges like land tenure insecurity, insufficient participant incentives, and persistent drivers of deforestation (illegal logging, large-scale concessions). Finally, an assessment of economic inequality and polarization associated with the Laos rubber boom demonstrates the importance of assessing how VSPs influence economic inequality. It also indicates that VSPs must address inequality's systemic drivers – including dispossession from land and forest resources, lacking worker protections, livelihood vulnerability, and barriers for smallholders – to maximize potential benefits. Overall, this dissertation research provides an example of how evidence synthesis, quasi-experimental methods, and consideration of economic, social, *and* environmental sustainability can deepen our understanding of VSPs.

EFFECTIVELY EVALUATING ENVIRONMENTAL, SOCIAL, AND ECONOMIC OUTCOMES IN VOLUNTARY SUSTAINABILITY PROGRAMS: LESSONS FROM LAOS

by

Rebecca Richter Traldi

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park, in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2022

Advisory Committee:
Associate Professor Julie A. Silva, Chair
Professor George C. Hurtt
Professor Michael Paolisso
Associate Professor Peter M. Steiner
Assistant Research Professor Alexandra Tyukavina

Dedication

To my mother, Sharon

Acknowledgements

I would like to thank the following people whose support was integral to the completion of my dissertation,

Johan, Maisy, Ruby, and Bo, thank you for the unconditional love and support. All my love;

To my family and friends – Mom, Dad, Oliver, Matthew, and Arthur, Phyllis, Malcolm, and Ila, the Occidental crew, Alix, Erin, and my Panda family, ARAGOG friends – thanks for always lending an ear when I needed it and believing in me;

Nonno, Evie, and Julie;

Allison, your support means everything; I continue to be dazzled by your intelligence, hilarity, and kindness;

Ipsita, thanks for the Whatsapps and for always being there for me;

Jeff, thank you for sharing your collaboration, knowledge, and friendship;

Thank you to the Geog PhD community, Dr. Rachel Berndston, and Dr. Kelly Anderson;

Ronnie, Ben, and Danny, thanks for all the laughs;

The community at the Thoracic Outlet Syndrome support group, thank you for the kind words, empathy, knowledge, validation, and brainstorming on solutions;

The community at the Disabled Academic Collective, thank you for hearing me;

My medical team, including Neema, Dr. Itskevitch, Dr. Lum, Dr. Tinsley, Dr. Bayat, Dr. Robinson, Nancy, Brigitte, Dr. Akman, Dr. Mai, and all of their nurses and support staff;

All of the scholars whose work helped inform and inspire this dissertation;

Everyone who has ever asked a question or contributed an answer on Stack Overflow;

Rya Inman, Johan Holmberg, Ashley Peña, Allison Baer, and Dragon dictation software for support formatting this dissertation;

Dan Newsome at the Center for Adaptive Technology at UMD;

Dr. Potapov, Dr. Tyukavina, and all of the collaborators in the Shifting Cultivation at a Crossroads Project, without whom this work would not have been possible;

I greatly appreciate funding provided by NASA-LCLUC (award #80NSSC20K0363), the UMD-Geography Human Dimensions of Global Change Summer Research Fellowship, and the BSOS-Dean's Research Initiative Fellowship.

Dr. Silva and my committee, I appreciate you investing your expertise and attention into my work and professional development during an incredibly challenging time.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

4C – Common Code for the Coffee Community

AR6 – Sixth Assessment Report

ARD –Analysis Ready Data

ATT – Average Treatment Effect on the Treated

BACI – Before and after control impact

CDE – Center for Development and Environment of the University of Bern

COSA – Committee for Sustainability Assessment

DAGs – Directed Acyclic Graphs

DER – Duclos Esteban Ray Index

FAO – Food and Agriculture Organization of the United Nations

FPIC – Free, Prior, Informed Consent (FPIC)

FSC – Forest Stewardship Council

GAP – Global Good Agricultural Practices

GPSNR – Global Platform on Sustainable Natural Rubber

IFOAM – Organic cropland

ILO - International Labour Organization

IPCC – Intergovernmental Panel on Climate Change

IPTW – Inverse Probability of Treatment Weighting

IPWR - Inverse probability-of-treatment weighted regression

IRSG – International Rubber Study Group

LCLUC – Land cover land-use change

LECS – Laos Expenditure and Consumption Survey (LECS)

LFA – Land and Forest Allocation Policy

LSB – Lao Statistics Bureau

MAF – Ministry of Agriculture and Forestry

NTFP – Non-Timber Forest Product

PEFC – Programme for the Endorsement of Forest Certification

SDGs – Sustainable Development Goals

SDPs – Sustainable Development Projects

SNRI – Sustainable Natural Rubber Initiative

VSPs – Voluntary Sustainability Programs

VSSs – Voluntary Sustainability Standards

WFP – World Food Programme

Chapter 1: Introduction

1.1 Motivation of the Study

In March 2022, the Intergovernmental Panel on Climate Change (IPCC) released its Sixth Assessment Report (AR6) highlighting the unprecedented risks of global climate change, and the urgency of mitigation and adaptation. But the report also notes that mitigation efforts – including afforestation and habitat restoration – may bring about unintended consequences, including increased vulnerability and resource competition (IPCC 2022). AR6 illustrates an important point for conservation practice – efforts to address environmental problems necessitate trade-offs and may engender negative outcomes. The IPCC's warning also underscores why rigorous impact evaluations of sustainability programs are crucial. For voluntary sustainability programs (VSPs), outcomes also affect program participation, since people are unlikely to continue engaging in programs which bring them no tangible benefits or introduce additional challenges (Pinto and McDermott 2013, Dayer et al. 2018). However, impact evaluations of VSPs – which by definition require a counterfactual assessment of what outcomes may have transpired in absence of the intervention – have historically been rare (Mascia et al. 2014, Fisher et al. 2014, Milder et al. 2015, Puri et al. 2017). Thus, this dissertation is motivated by the need for more frequent and careful evaluation of VSPs to demonstrate their environmental, economic, and social impacts and enable continuous improvement.

Since the 1990s, VSPs have become increasingly popular, utilized by a wide range of stakeholders, including governments, the private sector, and civil society organizations (Borck and Coglianese 2009). VSPs include voluntary sustainability standards (VSSs) like Rainforest Alliance and Fairtrade, and sustainable development projects which aim to improve rural livelihoods and natural resource management outcomes at the community level (Agrawal, Wollenberg, and Persha 2014, Hajjar et al. 2021, Potts et al. 2014). The former focus on the farm or co-op level certification of goods and rely primarily on market incentives, while the latter represent a wider range of intervention models, including participatory resource management and forest conservation. These types of interventions are similar in that they both provide incentives for participants to work towards improved sustainability outcomes, typically aiming to address all three pillars of sustainability (environmental, economic, and social), although with varying levels of priority. They usually provide resources to participants, whether through training, extension services, or other activities. The interventions' outcomes are meant to directly benefit participants, despite opportunity costs that might arise from joining. For example, agricultural certification targets socio-economic benefits like improved productivity, higher selling prices, or improved knowledge of agricultural best management practices (Potts et al. 2014). Sustainable development projects like community forestry often promote social sustainability, including livelihood security and poverty reduction (Santika et al. 2019).

Figure 1.1 illustrates different types of VSPs. In this dissertation, Chapter 2 focuses on VSSs; Chapter 3 focuses predominantly on sustainable agriculture (specifically, food security and livelihoods projects) and sustainable forest management; and Chapter 4 focuses on both VSSs and sustainable forest management programs.

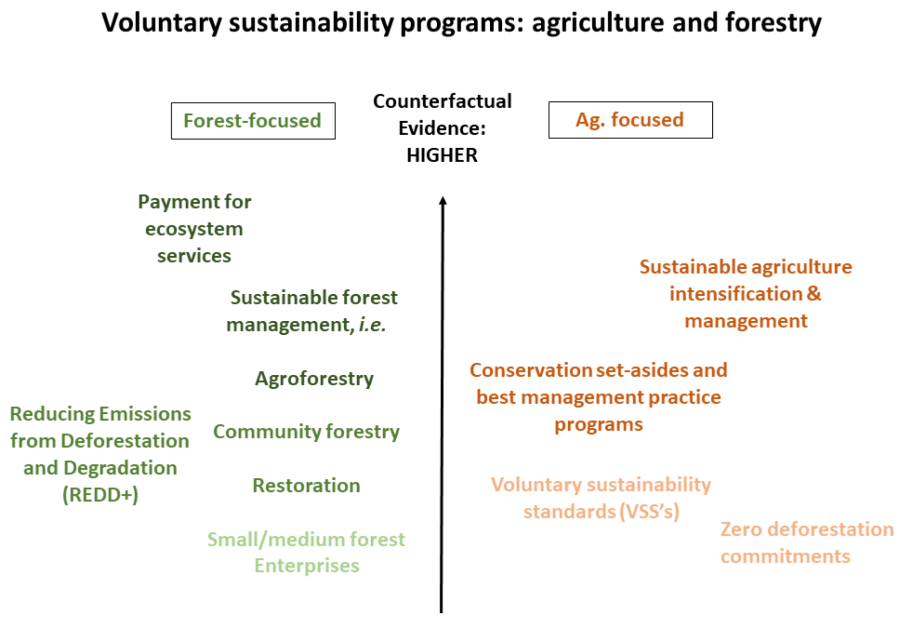


Figure 1.1: Illustration of a selection of voluntary sustainability programs (VSPs) focused on agriculture and forestry. The amount of existing counterfactual evidence regarding VSP impact increases from the bottom to the top of the figure, and with increasing color intensity. Source: Author's elaboration, drawing upon Agrawal, Wollenberg, and Persha 2014, Hajjar et al. 2021, Puri et al. 2017.

Calls for more robust evidence regarding the impact of VSPs are not new. But attention to this topic has historically been limited, despite research clearly outlining the value of impact evaluation for environmental interventions (Blackman and Rivera 2010, Ferraro and Hanauer 2014, Mascia et al. 2014). This is due to several factors, including the resource-intensive nature of monitoring and evaluation, inconsistent methods for assessing impact, fear of negative results, and the practical and ethical implications of analyzing outcomes in complex human-environment systems, which exhibit different pre-existing characteristics and enabling conditions (Baylis et al. 2016, Pressey et al. 2021). Additionally, studies have tended to focus on one aspect (or pillar) of sustainability, rather than considering environmental, economic, and social factors simultaneously (DeFries et al. 2017). Geography as a discipline is uniquely placed to inform evaluation of VSP impact for multiple sustainability pillars, since it enables the spatially and temporally explicit synthesis of environmental, economic, and social data (Bennett et al. 2017, Turner, Lambin, and Reenberg 2007).

One particularly important topic in the evaluation of VSP outcomes is that of trade-offs. For decades, sustainability science has emphasized the interdependencies between human well-being and environmental health (Spangenberg 2011). Sustainable development, at its core, is a vision for improving the human condition while sustaining the earth's resources (World Commission on Environment and Development 1987). But in real-world implementation, trade-offs between human development and environmental goals are common. Improved resource management may necessitate foregone economic opportunities; conversely, a focus on short-term economic productivity may disadvantage longer-term environmental sustainability (Vanderhaegen et al. 2018, Delgado-Serrano 2017, Burivalova et al. 2017). In addition, there is strong evidence that climate change will have the most severe impact in poor communities, worsening current poverty and inequality, with most of the costs being borne by developing countries (IPCC 2021).

Although there are multiple empirical methods to evaluate VSP impact, one critical approach is the estimation of counterfactuals. Essentially, a counterfactual outcome is the outcome that would have transpired with no intervention (Ferraro and Hanauer 2014). Counterfactual approaches are distinct from other methodologies in a few important ways. Namely, they eliminate rival explanations for results through either experimentation, conditioning (e.g. via matching, regression, fixed effects), or natural experiments (e.g. through instrumental variable approaches, Ferraro and Hanauer 2014). They also necessitate the consideration of causal assumptions. Therefore, they represent arguably the most robust approach to assessing the impact of VSPs. In this dissertation, I use the term 'counterfactual approach' interchangeably with 'impact evaluation' following Mascia et al. 2014 (see Table 1.1). Quasi-experimental research designs are a subset of counterfactual approaches which rely on observational rather than experimental data to explore causal questions (Ferraro and Hanauer 2014, Butsic et al. 2017, Jones and Lewis 2015).

Table 1.1: Summary of focal questions for five approaches to conservation monitoring and evaluation. Adapted from Mascia et al 2014. This research focuses on impact evaluation and systematic review.

Ambient Monitoring	**Management assessment**	**Performance measurement**	**Impact evaluation**	**Systematic review**
What is the state of social and/or environmental conditions, and how are they changing over time and space?	What are the management inputs, activities, and outputs […] and how are these changing over time?	To what extent is [the] intervention making progress toward its intended objectives for activities, outputs, and outcomes?	What intended and unintended impacts are causally induced by a conservation intervention?	What is the state of the evidence for the impact of an intervention and what does this evidence say about intervention impacts?

Although counterfactual analyses assessing VSPs are notoriously sparse, the last few years have seen an uptick in this research (Desbureaux 2021). As the field grows, disagreements have persisted within the conservation community about how progress should be measured (Pressey et al. 2021). Additionally, there are cases when counterfactual studies have generated different results for the same interventions (Heilmayr et al. 2020, Ermgassen et al. 2020). Therefore, it is the perfect time for critical discussion and inquiry regarding how studies should be structured. Which outcome indicators should be assessed? Which methods are most appropriate to address confounding bias? Where, and for which interventions, should research be conducted?

In this dissertation, I first assess what previous research has been conducted on VSPs using counterfactual approaches, identifying best practices and gaps in our current understanding of these initiatives. I then demonstrate how a counterfactual approach can be used to assess the impact of VSPs on environmental and socioeconomic outcomes. Finally, I explore the extent to which evaluations should incorporate measures of economic inequality in future impact assessments. I focus particularly on agriculture and forestry VSPs due to their importance for ecosystem health and sustainable development (Dudley and Alexander 2017, Kumar and Yashiro 2014, IPCC 2021). I juxtapose one global-scale analysis with two case studies of VSPs in Lao People's Democratic Republic (hereafter Laos).

Laos presents a useful case study for evaluating the impact of VSPs. Poverty and environmental challenges are both prevalent in Laos (Phompila et al. 2017, Bader et al. 2017). Deforestation and forest degradation are significant concerns, driven by a variety of factors including logging, mining, hydropower, large-scale plantations, and shifting cultivation (World Bank Group and Climate Investment Funds 2019, Zeng et al. 2018, Boutthavong, Hyakumura, and Ehara 2017). This has prompted a variety of actions from policymakers, NGOs, and other stakeholders, including bans on logging and large-scale plantation establishment from the government, as well as voluntary programs (World Bank Group and Climate Investment Funds 2019).

VSPs in Laos encompass a variety of foci, including food security and livelihoods, sustainable forestry production, and resource management projects in protected areas (see Appendix 2, section A2.2 and A2.3). These projects often address multiple pillars of sustainability with explicit goals of poverty reduction and livelihoods improvement, and sometimes include participatory components, including community involvement in identifying project objectives, strategies, and solutions (Williams et al. 2016, Boutthavong, Hyakumura, and Ehara 2017, Newby et al. 2012). Further, many VSPs in Laos assume that fostering livelihood security will help stem negative environmental impacts, including deforestation. This is due to the dominant narrative of small-scale shifting cultivation as a driver of forest loss and degradation, and the hypothesis that livelihoods improvements will disincentivize this practice and favor permanent agriculture and other forms of livelihood diversification (World Bank Group and Climate Investment Funds 2019). While there have been several published case studies exploring the deployment of VSPs in Laos, there has not been a national-level assessment of the outcomes of these initiatives.

1.2 Research design and dissertation structure

The research presented in this dissertation takes a mixed methods approach to assess the outcomes of VSPs for sustainable agriculture and forestry and consists of three main components. The second chapter is a systematic review of 45 studies evaluating the impact of global agricultural VSSs - a specific type of VSP. This informs the research design for two analyses using Laos as a case study. The third chapter analyzes a specific set of VSPs in Laos from 2005 to 2015, utilizing inverse probability-of-treatment weighted regression informed by directed acyclic graphs. This chapter assesses 72 projects in 1,452 villages, with a control sample size of 6,615. Then, the fourth chapter considers the sustainability outcomes of economic inequality and polarization, based on a case study of the Laos rubber boom from 2000-2011 (Household n size=4,924).

This study's findings offer important insights regarding the future direction of impact evaluation in conservation and sustainability programs, demonstrating appropriate methodological approaches for evaluating impacts and the specific contexts in which this assessment remains most critical. All three empirical chapters address the linkages between environmental and socioeconomic sustainability goals in VSPs. Chapter 2 clearly identifies research gaps as well as the most prominent methods previously utilized in quasi-experimental evaluations. It thus informs the methodological design for Chapter 3, which employs a counterfactual approach to explore VSP impact, specifically through a statistical matching design: inverse probability-of-treatment weighted regression to calculate the average treatment effect on the treated (ATT). Then, Chapter 4 takes a more forward-looking perspective, exploring the case for a higher prioritization of economic inequality as a sustainability outcome in the design and assessment of forest management VSPs, specifically related to rubber production. Figure 1.2 illustrates key features of Chapters 2-4. Finally, Chapter 5 outlines the study's conclusions and emphasizes main research findings, key contributions, and policy and research recommendations. Appendices are presented separately for each chapter and include additional elaboration on methods and supplementary results (A1-A3).

Key data sources utilized in this dissertation include (1) a database of 75 VSPs in Laos from 2005 to 2015 provided by the Center for Development and Environment of the University of Bern; (2) national estimates of LCLUC dynamics related to forest cover and rubber production area derived from Landsat time series data and calculated through supervised classification; (3) nationally representative poverty data from 2005 & 2015 Laos Population And Housing Census; and (4) nationally representative expenditures data from the 2007/08 & 2012/13 Laos Expenditures and Consumption Survey[1] (Lao DECIDE Info Project 2020, Potapov et al. 2019, 2020, Turubanova et al. 2018). Additionally, Chapter 3 involved extensive data curation and processing for a wide array of environmental, social, and economic covariates derived from publicly available data sources (see Appendix 2, Section A2.2). Finally, through Chapter 2, I created a database of previous impact evaluations of VSSs and made it publicly available upon publication[2].

[1] LECS data were provided by the Laos Statistics Bureau as part of the "Knowledge for Development" project led by the Center for Development and Environment of the University of Bern.
[2] The database is available as part of the Supplementary Materials for Chapter 2, via https://doi.org/10.1016/j.ecolind.2021.107490

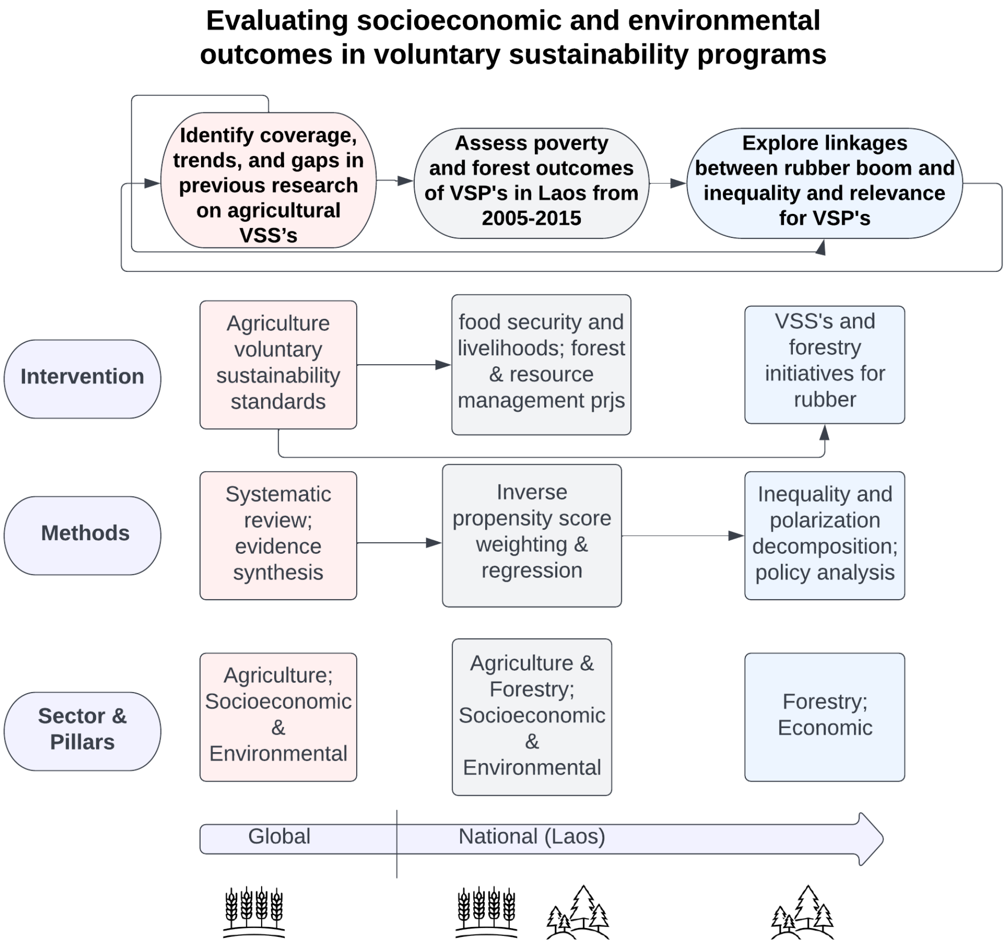


Figure 1.2: Overview of the study design, including research goals, interventions evaluated, and methods, as well as the sectors and sustainability pillars addressed in each chapter.

The study's research questions address three thematic sub-groupings within the broader topic of VSP impact. The research is organized in the following structure:

1.2.1 Chapter 2: Progress and Pitfalls: A systematic review of the evidence for agricultural sustainability standards

- Question 1a) Where are agricultural voluntary sustainability standards being evaluated through counterfactual approaches, and how does that compare to the extent of certification globally?

- Question 1b) Which pillars of sustainability are included in evaluation studies, and which indicators are used to measure success?
- Question 1c) What does the current evidence base suggest regarding outcomes of voluntary agricultural sustainability standards?

Questions 1a-1c relate specifically to global agricultural voluntary sustainability standards, which are typically deployed at farm level. This chapter distills information about the literature on impacts of agricultural VSSs to date, including the types of indicators evaluated and methods used. It builds on previous impact evaluations of VSSs (DeFries et al. 2017, Oya, Schaefer, and Skalidou 2018, Komives et al. 2018) and employs evidence synthesis methods to combine geospatial and statistical data, identifying trends and gaps. The methods follow the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines (Moher et al. 2009) and include a structured process for article search, selection, coding, and synthesis to draw conclusions regarding the first set of research questions.

By identifying specific crops, certifications, and geographies where evidence is most scarce (Question 1a), I clearly articulate a research agenda for the future evaluation of agricultural VSSs. The studies identified through this literature review also help inform the methods utilized in Question 2a, below. Additionally, Question 1b assisted in defining the research topic for Questions 3a and 3b, since the collation of previously evaluated indicators reveals that VSPs' influence on economic inequality is neglected in the literature.

1.2.2 Chapter 3: The poverty and forest cover impacts of sustainable development projects in Lao PDR from 2005 to 2015

- Question 2a) To what extent did voluntary sustainable development projects in Laos affect poverty and forest cover between 2005-2015?

This question scales down to the case study level, assessing village-level forest cover and poverty impacts of voluntary programs in Laos. Evaluating both socioeconomic and environmental outcomes simultaneously is a response to the research gaps identified by the systematic review in Chapter 2. The types of VSPs included in the study are varied, and include food security and nutrition, agricultural extension, sustainable forest management, agro-biodiversity, and poverty alleviation. I define a VSP project typology and assess poverty and forest cover outcomes separately for (1) all VSP projects, (2) food security and livelihoods projects, (3) forest management projects, and (4) resource management projects, thus exploring effect heterogeneity among different project types. I also create directed acyclic graphs (DAGs) to guide model design and the selection of 32 covariates to incorporate in the analysis (Elwert 2013, Rohrer 2018). The DAGs are based on content analysis of VSP program documents and literature review, and articulate important variables influencing selection into VSPs, and the relationship of confounders and VSPs to sustainability outcomes. The findings enable theoretical exploration of drivers of VSP effectiveness, juxtaposed with previous literature.

1.2.3 Chapter 4: Cultivating Inequality? Regional rubber dynamics and implications for voluntary sustainability programs in Lao PDR

- 3a) To what extent was rubber production associated with changes in rural economic inequality and polarization in Laos from 2007/08-2012/13?
- 3b) What are the implications of these trends for voluntary rubber sustainability programs?

Questions 3a and 3b focus on a case study of the Laos rubber boom. Examining the potential inequality-enhancing effects of rubber production is timely due to the recent proliferation of corporate commitments (e.g. from Michelin and Pirelli) in this sector, prompted by deforestation and other negative impacts during previous spikes in rubber production. In some ways, these questions flip the question of VSP impact on its head. Instead of evaluating the impact of historical VSPs, I ask how and why VSPs should consider specific sustainability outcomes - economic inequality and polarization - in the future. I explore these questions quantitatively through Gini and Duclos Esteban Ray index decomposition (Silva, Matyas, and Cunguara 2015, Araar 2006, Duclos, Esteban, and Ray 2004), with households as the primary unit of analysis. I combine socioeconomic data reflecting household expenditures with rubber production area estimates derived from Landsat imagery. I supplement my findings with an extensive literature review regarding socioeconomic impacts of the rubber boom in Laos, as well as policy analysis for VSPs. I explore the extent to which VSPs have considered economic inequality in the past and put forth policy recommendations based on my findings. I also discuss the sustainability implications of divergent agricultural and forestry production systems, illustrating the context-dependence of VSPs.

1.3 Bibliography

Agrawal, Arun, E. Wollenberg, and L. Persha. 2014. “Governing Agriculture-Forest Landscapes to Achieve Climate Change Mitigation.” *Global Environmental Change* 29 (November): 270–80. https://doi.org/10.1016/j.gloenvcha.2014.10.001.

Araar, Abdelkrim. 2006. “On the Decomposition of the Gini Coefficient: An Exact Approach, with an Illustration Using Cameroonian Data.” SSRN Scholarly Paper ID 877147. Rochester, NY: Social Science Research Network. https://doi.org/10.2139/ssrn.877147.

Bader, Christoph, Sabin Bieri, Urs Wiesmann, and Andreas Heinimann. 2017. “Is Economic Growth Increasing Disparities? A Multidimensional Analysis of Poverty in the Lao PDR between 2003 and 2013.” *The Journal of Development Studies* 53 (12): 2067–85. https://doi.org/10.1080/00220388.2016.1251587.

Baylis, K., J. Honey-Rosés, J. Börner, E. Corbera, D. Ezzine-de-Blas, P. J. Ferraro, R. Lapeyre, U. M. Persson, A. Pfaff, and S. Wunder. 2016. “Mainstreaming Impact Evaluation in Nature Conservation.” *Conservation Letters* 9 (1). https://doi.org/10.1111/conl.12180.

Bennett, Nathan J., Robin Roth, Sarah C. Klain, Kai Chan, Patrick Christie, Douglas A. Clark, Georgina Cullman, et al. 2017. “Conservation Social Science: Understanding and Integrating Human Dimensions to Improve Conservation.” *Biological Conservation* 205 (January): 93–108. https://doi.org/10.1016/j.biocon.2016.10.006.

Blackman, Allen, and Jorge E. Rivera. 2010. “The Evidence Base for Environmental and Socioeconomic Impacts of ‘Sustainable’ Certification.” *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.1579083.

Borck, Jonathan C., and Cary Coglianese. 2009. “Voluntary Environmental Programs: Assessing Their Effectiveness.” *Annual Review of Environment and Resources* 34 (1): 305–24. https://doi.org/10.1146/annurev.environ.032908.091450.

Boutthavong, Saykham, Kimihiko Hyakumura, and Makoto Ehara. 2017. “Stakeholder Participation in REDD+ Readiness Activities for Three Collaborative Projects in Lao PDR.” *Forests* 8 (5): 150. https://doi.org/10.3390/f8050150.

Burivalova, Zuzana, Fangyuan Hua, Lian Pin Koh, Claude Garcia, and Francis Putz. 2017. “A Critical Comparison of Conventional, Certified, and Community Management of Tropical Forests for Timber in Terms of Environmental, Economic, and Social Variables.” *Conservation Letters* 10 (1): 4–14. https://doi.org/10.1111/conl.12244.

Butsic, Van, David J. Lewis, Volker C. Radeloff, Matthias Baumann, and Tobias Kuemmerle. 2017. “Quasi-Experimental Methods Enable Stronger Inferences from Observational Data in Ecology.” *Basic and Applied Ecology* 19 (March): 1–10. https://doi.org/10.1016/j.baae.2017.01.005.

Dayer, Ashley A., Seth H. Lutter, Kristin A. Sesser, Catherine M. Hickey, and Thomas Gardali. 2018. “Private Landowner Conservation Behavior Following Participation in Voluntary Incentive Programs: Recommendations to Facilitate Behavioral Persistence.” *Conservation Letters* 11 (2): e12394. https://doi.org/10.1111/conl.12394.

DeFries, Ruth S., Jessica Fanzo, Pinki Mondal, Roseline Remans, and Stephen A. Wood. 2017. “Is Voluntary Certification of Tropical Agricultural Commodities Achieving Sustainability Goals for Small-Scale Producers? A Review of the Evidence.” *Environmental Research Letters* 12 (3): 033001. https://doi.org/10.1088/1748-9326/aa625e.

Delgado-Serrano, María Mar. 2017. "Trade-Offs between Conservation and Development in Community-Based Management Initiatives." *International Journal of the Commons* 11 (2): 969–91. https://doi.org/10.18352/ijc.792.

Desbureaux, Sébastien. 2021. "Subjective Modeling Choices and the Robustness of Impact Evaluations in Conservation Science." *Conservation Biology* 35 (5): 1615–26. https://doi.org/10.1111/cobi.13728.

Duclos, Jean-Yves, Joan Esteban, and Debraj Ray. 2004. "Polarization: Concepts, Measurement, Estimation." *Econometrica* 72 (6): 1737–72. https://doi.org/10.1111/j.1468-0262.2004.00552.x.

Dudley, Nigel, and Sasha Alexander. 2017. "Agriculture and Biodiversity: A Review." *Biodiversity* 18 (2–3): 45–49. https://doi.org/10.1080/14888386.2017.1351892.

Elwert, Felix. 2013. "Graphical Causal Models." In *Handbook of Causal Analysis for Social Research*, edited by Stephen L. Morgan, 245–73. Handbooks of Sociology and Social Research. Dordrecht: Springer Netherlands. https://doi.org/10.1007/978-94-007-6094-3_13.

Ermgassen, Erasmus K. H. J. zu, Ben Ayre, Javier Godar, Mairon G. Bastos Lima, Simone Bauch, Rachael Garrett, Jonathan Green, et al. 2020. "Using Supply Chain Data to Monitor Zero Deforestation Commitments: An Assessment of Progress in the Brazilian Soy Sector." *Environmental Research Letters* 15 (3): 035003. https://doi.org/10.1088/1748-9326/ab6497.

Ferraro, Paul J., and Merlin M. Hanauer. 2014. "Advances in Measuring the Environmental and Social Impacts of Environmental Programs." *Annual Review of Environment and Resources* 39 (1): 495–517. https://doi.org/10.1146/annurev-environ-101813-013230.

Fisher, Brendan, Andrew Balmford, Paul J. Ferraro, Louise Glew, Michael Mascia, Robin Naidoo, and Taylor H. Ricketts. 2014. "Moving Rio Forward and Avoiding 10 More Years with Little Evidence for Effective Conservation Policy: Making the Rio Declaration Work." *Conservation Biology* 28 (3): 880–82. https://doi.org/10.1111/cobi.12221.

FSC United States. 2020. "FSC Certificate Database." FSC United States. 2020. https://us.fsc.org:443/en-us/market/find-products/fsc-certificate-database.

Hajjar, Reem, Peter Newton, Markus Ihalainen, Arun Agrawal, Jennifer Alix-Garcia, Sarah E. Castle, James T. Erbaugh, et al. 2021. "Levers for Alleviating Poverty in Forests." *Forest Policy and Economics* 132 (November): 102589. https://doi.org/10.1016/j.forpol.2021.102589.

Heilmayr, Robert, Lisa L. Rausch, Jacob Munger, and Holly K. Gibbs. 2020. "Brazil's Amazon Soy Moratorium Reduced Deforestation." *Nature Food* 1 (12): 801–10. https://doi.org/10.1038/s43016-020-00194-5.

IPCC, 2022: Summary for Policymakers [H.-O. Pörtner, D.C. Roberts, E.S. Poloczanska, K. Mintenbeck, M. Tignor, A. Alegría, M. Craig, S. Langsdorf, S. Löschke, V. Möller, A. Okem (eds.)]. In: Climate Change 2022: Impacts, Adaptation, and Vulnerability. Contribution of Working Group II to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change [H.-O. Pörtner, D.C. Roberts, M. Tignor, E.S. Poloczanska, K. Mintenbeck, A. Alegría, M. Craig, S. Langsdorf, S. Löschke, V. Möller, A. Okem, B. Rama (eds.)]. Cambridge University Press. In Press. Jones, Kelly W., and David J. Lewis. 2015. "Estimating the Counterfactual Impact of Conservation Programs on Land Cover Outcomes: The Role of Matching and Panel Regression Techniques." *PLoS ONE* 10 (10). https://doi.org/10.1371/journal.pone.0141380.

Komives, Kristin, Ashleigh Arton, Ellen Baker, Elizabeth Kennedy, Catherine Longo, Deanna Newsom, Alexander Pfaff, and Claudia Romero. 2018. "Conservation Impacts of Voluntary Sustainability Standards: How Has Our Understanding Changed since the 2012 Publication of 'Toward Sustainability: The Roles and Limitations of Certification'?" Washington D.C.: Meridian Institute.

Kumar, Pushpam, and Makiko Yashiro. 2014. "The Marginal Poor and Their Dependence on Ecosystem Services: Evidence from South Asia and Sub-Saharan Africa." In *Marginality: Addressing the Nexus of Poverty, Exclusion and Ecology*, edited by Joachim von Braun and Franz W. Gatzweiler, 169–80. Dordrecht: Springer Netherlands. https://doi.org/10.1007/978-94-007-7061-4_11.

Lao DECIDE Info Project. 2020. "Lao DECIDE Info - Start." Lao DECIDE Info. 2020. http://www.decide.la/en/.

Mascia, Michael B., Sharon Pailler, Michele L. Thieme, Andy Rowe, Madeleine C. Bottrill, Finn Danielsen, Jonas Geldmann, Robin Naidoo, Andrew S. Pullin, and Neil D. Burgess. 2014. "Commonalities and Complementarities among Approaches to Conservation Monitoring and Evaluation." *Biological Conservation* 169 (January): 258–67. https://doi.org/10.1016/j.biocon.2013.11.017.

Milder, Jeffrey C., Margaret Arbuthnot, Allen Blackman, Sharon E. Brooks, Daniele Giovannucci, Lee Gross, Elizabeth T. Kennedy, et al. 2015. "An Agenda for Assessing and Improving Conservation Impacts of Sustainability Standards in Tropical Agriculture." *Conservation Biology* 29 (2): 309–20. https://doi.org/10.1111/cobi.12411.

Moher, David, Alessandro Liberati, Jennifer Tetzlaff, Douglas G. Altman, and PRISMA Group. 2009. "Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement." *PLoS Medicine* 6 (7): e1000097. https://doi.org/10.1371/journal.pmed.1000097.

Newby, Jonathan C., R. A. Cramb, Somphanh Sakanphet, and Sean McNamara. 2012. "Smallholder Teak and Agrarian Change in Northern Laos." *Small-Scale Forestry* 11 (1): 27–46. https://doi.org/10.1007/s11842-011-9167-x.

Oya, Carlos, Florian Schaefer, and Dafni Skalidou. 2018. "The Effectiveness of Agricultural Certification in Developing Countries: A Systematic Review." *World Development* 112 (December): 282–312. https://doi.org/10.1016/j.worlddev.2018.08.001.

Phompila, Chittana, Megan Lewis, Bertram Ostendorf, and Kenneth Clarke. 2017. "Forest Cover Changes in Lao Tropical Forests: Physical and Socio-Economic Factors Are the Most Important Drivers." *Land* 6 (2): 23. https://doi.org/10.3390/land6020023.

Pinto, Luís Fernando Guedes, and Constance McDermott. 2013. "Equity and Forest Certification — A Case Study in Brazil." *Forest Policy and Economics* 30 (May): 23–29. https://doi.org/10.1016/j.forpol.2013.03.002.

Potapov, Peter, Matthew C. Hansen, Indrani Kommareddy, Anil Kommareddy, Svetlana Turubanova, Amy Pickens, Bernard Adusei, Alexandra Tyukavina, and Qing Ying. 2020. "Landsat Analysis Ready Data for Global Land Cover and Land Cover Change Mapping." *Remote Sensing* 12 (3): 426. https://doi.org/10.3390/rs12030426.

Potapov, Peter, A. Tyukavina, S. Turubanova, Y. Talero, A. Hernandez-Serna, M. C. Hansen, D. Saah, et al. 2019. "Annual Continuous Fields of Woody Vegetation Structure in the Lower Mekong Region from 2000-2017 Landsat Time-Series." *Remote Sensing of Environment* 232 (October): 111278. https://doi.org/10.1016/j.rse.2019.111278.

Potts, Jason, Matthew Lynch, Ann Wilkings, Gabriel A Huppé, Maxine Cunningham, Vivek Anand Voora, ENTWINED (Organization), et al. 2014. *The State of*

Sustainability Initiatives Review 2014: Standards and the Green Economy. http://www.deslibris.ca/ID/242257.
Pressey, Robert L., Piero Visconti, Madeleine C. McKinnon, Georgina G. Gurney, Megan D. Barnes, Louise Glew, and Martine Maron. 2021. “The Mismeasure of Conservation.” *Trends in Ecology & Evolution* 36 (9): 808–21. https://doi.org/10.1016/j.tree.2021.06.008.
Puri, Jyotsna, Megha Nath, International Initiative for Impact Evaluation (3ie), Raag Bhatia, International Initiative for Impact Evaluation (3ie), Louise Glew, and World Wildlife Fund. 2017. “Examining the Evidence Base for Forest Conservation Interventions.” 2017th ed. International Initiative for Impact Evaluation (3ie). https://doi.org/10.23846/egm0004.
Rainforest Alliance. 2022. “Certified Products.” Rainforest Alliance. 2022. https://www.rainforest-alliance.org/find-certified.
Rohrer, Julia M. 2018. “Thinking Clearly About Correlations and Causation: Graphical Causal Models for Observational Data.” *Advances in Methods and Practices in Psychological Science* 1 (1): 27–42. https://doi.org/10.1177/2515245917745629.
Santika, Truly, Kerrie A. Wilson, Sugeng Budiharta, Ahmad Kusworo, Erik Meijaard, Elizabeth A. Law, Rachel Friedman, et al. 2019. “Heterogeneous Impacts of Community Forestry on Forest Conservation and Poverty Alleviation: Evidence from Indonesia.” *People and Nature* 1 (2): 204–19. https://doi.org/10.1002/pan3.25.
Silva, Julie A., Corene J. Matyas, and Benedito Cunguara. 2015. “Regional Inequality and Polarization in the Context of Concurrent Extreme Weather and Economic Shocks.” *Applied Geography*, Spatial Inequality, 61 (July): 105–16. https://doi.org/10.1016/j.apgeog.2015.01.015.
Smith, Hilary, Peter Kanowski, Rodney J. Keenan, and Somvang Phimmavong. 2021. “Lao Plantation Policy: Prospects for Change.” *Forests* 12 (8): 1132. https://doi.org/10.3390/f12081132.
Spangenberg, Joachim H. 2011. “Sustainability Science: A Review, an Analysis and Some Empirical Lessons.” *Environmental Conservation* 38 (3): 275–87. https://doi.org/10.1017/S0376892911000270.
Turner, B. L., Eric F. Lambin, and Anette Reenberg. 2007. “The Emergence of Land Change Science for Global Environmental Change and Sustainability.” *Proceedings of the National Academy of Sciences* 104 (52): 20666–71. https://doi.org/10.1073/pnas.0704119104.
Turubanova, Svetlana, Peter V. Potapov, Alexandra Tyukavina, and Matthew C. Hansen. 2018. “Ongoing Primary Forest Loss in Brazil, Democratic Republic of the Congo, and Indonesia.” *Environmental Research Letters* 13 (7): 074028. https://doi.org/10.1088/1748-9326/aacd1c.
Vanderhaegen, Koen, Kevin Teopista Akoyi, Wouter Dekoninck, Rudy Jocqué, Bart Muys, Bruno Verbist, and Miet Maertens. 2018. “Do Private Coffee Standards ‘Walk the Talk’ in Improving Socio-Economic and Environmental Sustainability?” *Global Environmental Change* 51 (July): 1–9. https://doi.org/10.1016/j.gloenvcha.2018.04.014.
Williams, Liana J., Sharmin Afroz, Peter R. Brown, Lytoua Chialue, Clemens M. Grünbühel, Tanya Jakimow, Iqbal Khan, et al. 2016. “Household Types as a Tool to Understand Adaptive Capacity: Case Studies from Cambodia, Lao PDR, Bangladesh and India.” *Climate and Development* 8 (5): 423–34. https://doi.org/10.1080/17565529.2015.1085362.
World Bank Group, and Climate Investment Funds. 2019. “A Case Study of the Lao PDR Forest Investment Program.”

http://documents1.worldbank.org/curated/en/701081563953173013/pdf/SUFORD-SU-Project.pdf.

World Commission on Environment and Development. 1987. *Our Common Future.* Oxford: Oxford University Press. https://sustainabledevelopment.un.org/content/documents/5987our-common-future.pdf.

Zeng, Zhenzhong, Lyndon Estes, Alan D. Ziegler, Anping Chen, Timothy Searchinger, Fangyuan Hua, Kaiyu Guan, Attachai Jintrawet, and Eric F. Wood. 2018. "Highland Cropland Expansion and Forest Loss in Southeast Asia in the Twenty-First Century." *Nature Geoscience* 11 (8): 556–62. https://doi.org/10.1038/s41561-018-0166-9.

Chapter 2: Progress and Pitfalls: a systematic review of the evidence for agricultural sustainability standards[3]

2.1 Introduction

Agricultural production is critical for humanity's survival. While agriculture provides an indispensable service, it also results in serious consequences for environmental and social sustainability outcomes. Agriculture is known to be a key driver of negative environmental impacts, including deforestation and subsequent impacts on wildlife habitat and greenhouse gas emissions, nutrient imbalances due to intensive fertilizer application and other practices, and impacts on soil and water resources (Foley et al., 2011, Vitousek et al., 2009). Production can also have a range of implications for human well-being, from changes in livelihoods activities to violations of labor regulations, child labor, and forced labor, among other issues (O'Rourke, 2014, Rasmussen et al., 2018).

A variety of interventions have been implemented to stem these negative externalities and promote resource stewardship and benefits for local communities. One significant category of interventions is voluntary sustainability standards, or VSSs. Voluntary sustainability standards first came to the forefront in the 1980s, with standards like Organic (IFOAM) and the Rainforest Alliance (Potts et al., 2014). Voluntary standards are based on the idea that a combination of positive incentives (e.g. price premiums for producers and/or provision of other services), training and awareness building, clear and consistent criteria for success, and a market-based approach can join forces to boost sustainability (Smith et al., 2019). Most voluntary standards outline requirements related to

[3] Note that this chapter has been published in Ecological Indicators. Please see the full citation: Traldi, Rebecca. 2021. "Progress and Pitfalls: A Systematic Review of the Evidence for Agricultural Sustainability Standards." *Ecological Indicators* 125 (June): 107490. https://doi.org/10.1016/j.ecolind.2021.107490.

social, economic, and environmental sustainability, although the specific principles, criteria and indicators vary among the standards, as illustrated by Table 2.1. There are hundreds of these standards globally, including in forestry and seafood sectors (IISD, 2015). For agriculture, there are about 13 standards that are the most widely adopted and recognized by the international community (Willer et al., 2019). According to recent research, agricultural standards represent about 1.1% of global agricultural area, although their production is not equally distributed among regions (Tayleur et al., 2017).

Table 2.1: Overview of 13 major international agricultural voluntary sustainability standards and their key characteristics, principles, and criteria[4]. There is overlap between sustainability pillars within principles and criteria. Data is from program websites, the ITC Standards Map (2015), Potts et al. 2014, and Willer et al. 2019.

Standard	**Crop specific?**	**Process-based vs. performance-based[5]**	**Year initiated**	**Environmental principles and criteria (P&C)**	**Social P&C**	**Economic P&C**
Better Cotton Initiative	Yes	Process-based	2005	Crop protection, water stewardship, soil health, biodiversity, responsible land use	Decent work conditions	Fiber quality, management systems
Bonsucro (sugarcane)	Yes	Performance-based	2008	Manage biodiversity and ecosystem services, additional biofuel requirements under EU renewable energy directive	Obey the law, respect human rights and labor standards	Production efficiency, continuously improve key areas of the business
Common Code for the Coffee Community (4C)	Yes	Process-based	2006	Biodiversity, energy, soil management, waste management, water management	Work and labor rights, working conditions, gender, health and safety	Profitability and productivity, capacity development, record keeping, market access/information, quality, traceability
Cotton Made in Africa	Yes	Process-based	2005	Responsible land use, enhance biodiversity, and	Responsible business conduct, support	Effective management systems; access to high

[4] This table is intended to provide an overview, rather than a comprehensive list of all standard criteria. C.A.F.E Practices and Bird Friendly certifications are not covered here. For a more detailed overview regarding the environmental coverage of these standards, see Tayleur et al. 2017.
[5] Process-based standards outline practices that must be implemented, but not specific outcomes; performance-based standards specify actual outcomes to be achieved. Both approaches can exist within one standard (Potts et al. 2014).

				protect climate and environment; GMO-free cotton, care for water and soil; minimize adverse impacts of crop protection	smallholder farmers, decent work, respect children's' rights and gender equality	quality inputs and pre-financing; increase productivity and fiber quality; improving living conditions and resilience
Fairtrade	No	Process-based	1997	Agricultural practices e.g. agrochemicals, waste, soil and water, GMOs	Social development, e.g. organizational transparency, worker rights and security, working conditions	Required minimum price and/or price premium (the latter is invested in quality of life improvements), pre-financing
Global Good Agricultural Practices (GAP)	No	Process-based	1997	Waste and pollution management, conservation	Worker health, safety, and welfare, complaints management	Site management, record-keeping, hygiene, recall procedure
Organic cropland (IFOAM)	No	Process-based	1972	Organic ecosystems, crop production	Social justice	Processing and handling
Proterra	No	Process-based	2012	Biodiversity conservation, effective env. management; no GMOs; pollution and waste mgmt.; water mgmt.; GHG and energy; adoption of good ag. practices	Compliance with law; human rights and responsible labor practices; responsible relations with workers & community	Traceability and chain of custody
Rainforest Alliance/Sustainable	No	Process-based; some outcome criteria (e.g.	1987	Biodiversity conservation, natural resource conservation	Improved livelihoods and human well-being	Effective planning and management

Agricultural Network		specific native vegetation thresholds)			(e.g. working conditions)	
Roundtable on Responsible Soy	Yes	Process-based	2006	Environmental responsibility, good agricultural practices	Legal compliance, responsible labor conditions & community relations	Good business practices
Roundtable on Sustainable Biomaterials[6]	Yes	Process-based	2007	GHG emissions, conservation, soil, water, air	Legality, human and labor rights, rural development, local food security, land rights	Planning, monitoring and continuous improvement; use of technology, inputs, and waste management
Roundtable on Sustainable Palm Oil	Yes	Process-based	2004	Protect, conserve, and enhance ecosystems and the environment	Behave ethically and transparently; operate legally; respect human rights; support smallholder inclusion; respect workers' rights and conditions	Optimize productivity, efficiency, positive impacts, and resilience
UTZ[7]	No	Process-based	2002	Soil, waste, water, biodiversity, energy	Labor rights, health and safety, employment conditions, human rights	Price premiums

[6] Certifies several crops, residues, and associated feedstocks, including sugar cane, waste starch from wheat, coconut, *brassica carinata*, jatropha and corn.
[7] UTZ merged with Rainforest Alliance in 2018. Since this is relatively recent, UTZ is treated as a separate standard in this review.

While there are slight variations among the standards, they generally operate in the following way: producers voluntarily commit to improving their sustainability practices by adhering to the principles and criteria of the standard (for example improving farm efficiency, implementing conservation measures, or ensuring social safeguards). They complete the transition with support from the standard and possibly other stakeholders, depending on the local context. Their operations are then audited by an impartial third party. Finally, their product is "certified", and may be labeled with the appropriate standard seal, depending on the certification (Milder et al., 2015).

Many different stakeholders participate in voluntary standards design and implementation, including non-governmental organizations, private companies, and industry associations. Standards are also a key sustainability tool for many large international brands who source significant amounts of agricultural raw materials. Companies like Coca-Cola, PepsiCo, and many others have committed to sourcing certified crops as part of their sustainability strategies (Smith et al., 2019).

2.1.1 The need for empirical evaluation of certification's impacts

Over the last decade, there has been increasing interest in understanding the efficacy of voluntary standards in achieving agricultural sustainability objectives. In 2011, Blackman and Rivera emphasized the limited number of robust studies of the impact of certification, arguing that the majority did not rely on research designs which could reasonably indicate causation. This was largely due to inadequate incorporation of a "counterfactual" approach, or a test of what may have happened in the absence of the certification (Blackman & Rivera, 2011). In practice, measuring the counterfactual outcome typically requires comparison between a “treatment” group that has been certified, and a “control” group that has not been certified (Blackman & Rivera, 2011).

Since then, there have been significant developments in assessing certification’s outcomes. Notably, there have been two systematic reviews focused on the impacts of agricultural certification in the last five years, published by Oya et al. (2018) and DeFries et al. (2017). The former focuses primarily on socioeconomic outcomes for producers, while the latter looks at outcomes across all pillars of sustainability. These reviews have enabled a clearer picture of the evidence regarding certification’s impacts, and set out structured protocols for identifying, reviewing, and selecting studies for inclusion. They also indicated a growing but still imbalanced evidence base, since over 80% of the studies included in DeFries et al. 2017 were focused on coffee certification. The primary statistical methods utilized in robust impact evaluations of certification include multivariate regression, matched pair comparisons, propensity score matching, instrumental variable approaches, and difference-in-difference methods (DeFries et al. 2017, Ferraro & Hanauer, 2014). Each of these methods necessitates consideration of various assumptions. For example, matching designs require that selection into the certification program occur only due to observed variables (Bolwig et al., 2009). The instrumental variable approach requires identification of an instrument that is correlated with the treatment variable, but not directly correlated with outcomes of interest (Chiputwa & Qaim, 2016). Broadly, these methods fall into the realm of “quasi-experimental” research designs (see Ferraro, 2009, and Butsic et al. 2017).

There is considerable interest in evaluating the impacts of certification from sustainable development practitioners, evidenced by a growing amount of grey literature. For example, many non-government organizations have published their own reports to distill insights on outcomes of certification (Komives et al., 2018, Petrokofsky & Jennings, 2018). In 2019, an online database called Evidensia was launched with the explicit goal of making certification impact studies widely available (evidensia.eco).

There are also parallels between empirical work evaluating VSS certification, and the evaluation of other sustainability interventions. These evaluations depend on the use of indicators which appropriately measure relevant sustainability outcomes, and are important for a wide range of disciplines from both a theoretical and applied perspective. For example, indicators have been used to assess best management practices and conservation on farmlands (Targetti et al., 2014, Last et al., 2014, Garibaldi et al., 2017, Latruffe et al., 2016), the socioeconomic and environmental impacts of protected areas (Naidoo et al., 2019), and the outcomes of ecosystem management (Breslow et al. 2016, Breslow et al. 2017).

Despite recent developments, substantial gaps remain in our understanding of the impact of agricultural certification. Given the severity of current sustainability challenges, stakeholders may wonder whether changes are needed to the certification model, or question whether and how these interventions should be integrated with other activities at the landscape level (Tscharntke et al., 2015). It's also unclear to what extent general lessons learned can be inferred from study results to date, since evaluations are not representative of all certifications, production systems, or regions globally.

This study responds to these persistent questions regarding the impacts of voluntary sustainability standards for agriculture at the plot, farm, and household level, building upon previous reviews. Specifically, it addresses the questions:

1) Where are certifications being studied, and how does that compare to the extent of certification globally?

2) Which pillars of sustainability are included in evaluation studies, and which indicators are used to measure success?

3) What does the current evidence base suggest regarding outcomes of voluntary agricultural sustainability standards?

The analysis complements the literature on impacts of agricultural VSSs in a few ways. First, over half of the studies assessed here have not been included in previous reviews, due to their recent publication. Since the empirical evaluation of VSSs is a rapidly developing field, we can now ask more nuanced questions regarding research gaps than was previously possible. Second, this analysis provides a new level of detail regarding the distribution of impact evaluations of certification by crop, certification, and country, and identifies specific research gaps and needs. Third, by blending aspects of quantitative and qualitative systematic review, this study synthesizes information about methods to measure success. This enables identification of best practices that can help inform consistent and clear outcome indicators in future impact evaluations. In addition to the literature regarding VSSs, these insights can be applied more broadly to studies measuring the sustainability outcomes of environmental interventions.

2.2 Materials and Methods

There were five main stages of the literature review approach. First, a search was conducted to identify articles to screen for inclusion in the analysis. Second, each article was assessed based on a predefined list of inclusion criteria, based on previous reviews (primarily DeFries et al. 2017 and Oya et al. 2018), and an initial list of articles for inclusion was prepared. Third, each of these studies was read and evaluated for key insights, including research design and main findings. Fourth, a detailed indicator table was developed for each study, following DeFries et al. 2017. Finally, findings were synthesized with other data sources to address the research questions and distill key areas of progress and gaps, and descriptive statistics were calculated. More information on each stage of the process is provided below.

This study provides a methodical, replicable, and transparent approach to collecting evidence on VSS outcomes. This is achieved by clearly outlining the search process, inclusion criteria, and evaluation method (Siddaway et al., 2019). A schematic representing the key methodological stages and following PRISMA guidelines (Moher et al. 2009) is shown in Figure 2.1. A PRISMA checklist can be found in the Appendix.

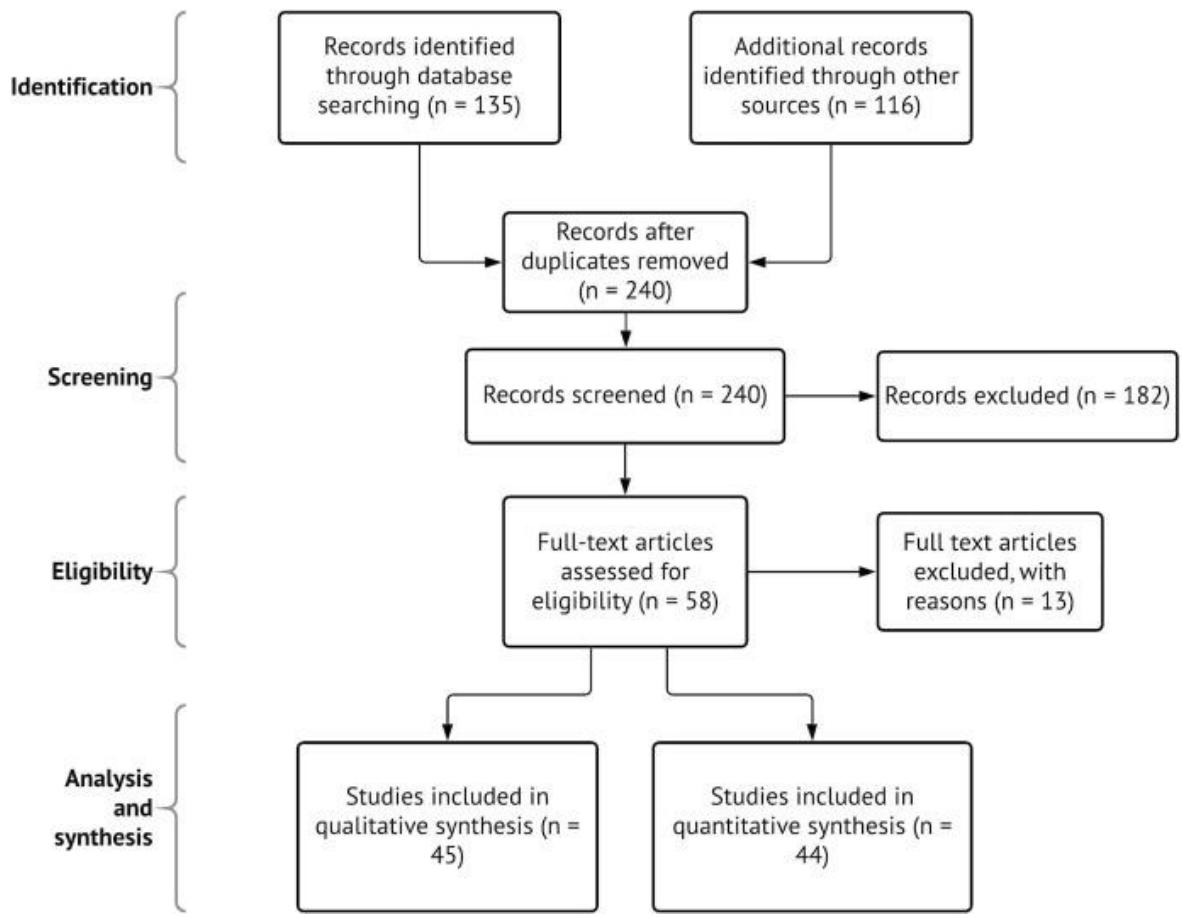


Figure 2.1: Methodology diagram, following PRISMA guidelines (Moher et al. 2009).

2.2.1 Identifying articles

Literature searches were conducted through the following sources: 1. Publications selected by recent reviews, 2. All agricultural articles housed in the Evidensia database as of November 2019, and standardized searches through 3. Google Scholar, and 4. Web of Science. The Web of Science search terms included (impact of sustainability certification AND agriculture*), which returned 135 articles as of January 2020. A small number of sources were identified through citations within research articles from the initial search. In total, 240 articles were screened - 135 identified through Web of Science, and 116 identified through other sources listed above. All article searches were conducted only in English, and only English-language articles are included in the analysis. The article list was finalized in May 2020.

2.2.2 Screening articles

The first round of screening relied primarily on the abstract. Articles which did not include a counterfactual, quasi-experimental research design to assess the impact of certification were eliminated. In cases where the abstract did not provide sufficient information to ascertain this, the article was read to determine the methodological approach and the extent to which selection bias and group equivalence between treatment and control groups were addressed. Articles that were previously characterized as having high bias (e.g. were noted as "Critical Bias" by Oya et al. 2018's assessment or designated as high bias by Komives et al. 2019) were not included in the analysis. Four articles were identified after the article list was finalized, and thus were not included (Dietz et al., 2018, Dietz et al., 2020, Sellare et al. 2020, Santika et al. 2020; the first two focus on coffee, the latter two on cocoa and palm oil respectively). For an overview of the reviewed articles, see Table 2.2. Among the included articles, 19 have been included in previous quantitative reviews, and 26 have not.

2.2.3 Eligibility, analysis, and synthesis

Each study was read and coded for various characteristics, including the region, country, certification, and crop studied, the research design and sample size, the pillars of sustainability included in the analysis, and key findings. Some initial characteristics that were appealing to extract from studies were not consistently available - for example, the years since certification was obtained for the study population. At this stage, a few articles initially selected for inclusion were eliminated for various reasons, e.g. methodological approach, resulting in an n size of 45 articles and 66 cases. Several studies include more than one case (multiple certifications in one country, or more than one country of analysis).

For each indicator, the pillar of sustainability and general theme was recorded, as well as the indicator itself, and the results, noting no difference, positive, and negative outcomes (see Table 2.3). If the significance level of an indicator was uncertain after reviewing a study's results tables, the corresponding author was contacted to clarify. As Figure 2.2 illustrates, outputs relate to direct activities resulting from an intervention, while outcomes refer to longer term goals including changes in behavior, social, and environmental conditions (Mascia et al. 2014). Indicators which were contextual in nature or unrelated to a certification outcome were classified as "Other" rather than being incorporated within a sustainability pillar. Overall, 777 indicators were included in the review (see online Supplementary Materials for the full indicator table).

To answer the research questions, summary statistics were calculated for key elements of the included studies, including indicator variables. These results were synthesized with other data sources - for example, for question 1, findings were synthesized with insights from Willer et al. 2019 and Tayleur et al. 2017 on the global extent of certification. Finally, indicators grouped by sustainability pillar and by theme (e.g. income, productivity, best management practices) were tested for significant differences using the Kruskal-Wallis analysis of variance in R (R Core Team 2013).

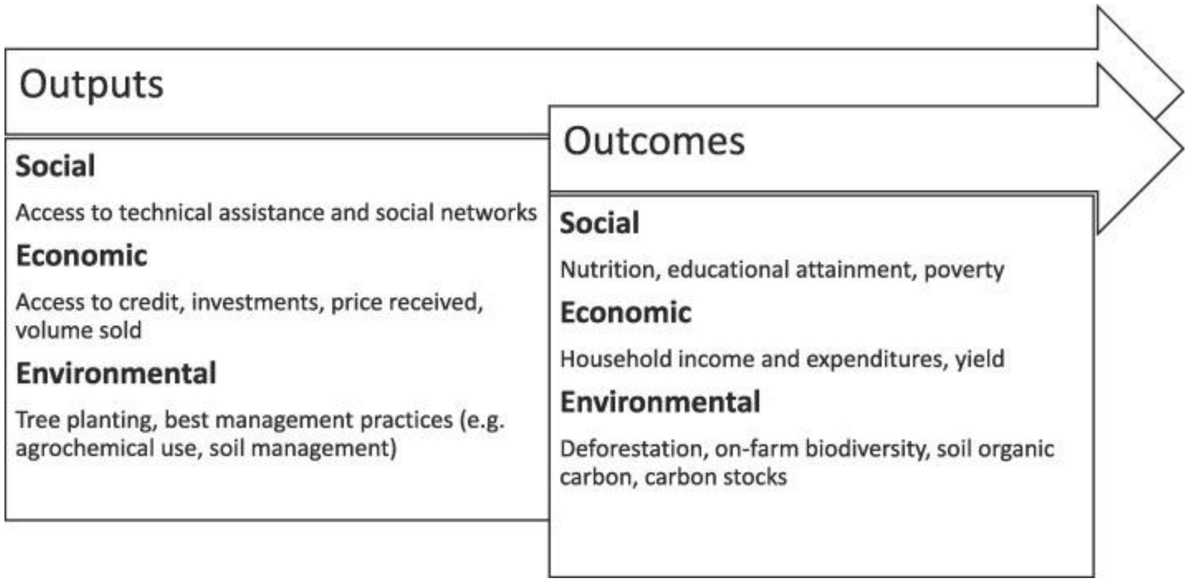


Figure 2.2: Outputs versus sustainability outcomes for voluntary sustainability standards. Standards' outcome goals vary; the figure presents illustrative examples.

Table 2.2: Studies included in this review, in order of publication year. Some studies include multiple cases or analyses of certification (e.g. multiple crops, countries, or certifications). Combined certification is indicated with a hyphen (e.g. "Fairtrade-Organic").

Study	**Location(s)**	**Crop**	**Certification(s)**	**Methods**	**Included in previous review[8]?**
Arnould et al., 2009	Nicaragua, Peru, and Guatemala	Coffee	Fairtrade	OLS and binomial logistic regression	Yes
Bolwig et al., 2009	Uganda	Coffee	Organic	OLS regression and maximum likelihood estimation	Yes
Fort & Ruben, 2009	Peru	Banana	Fairtrade	Propensity score matching, probit regression	Yes
Ruben et al., 2009	Costa Rica	Coffee, Banana	Fairtrade	Propensity score matching	Yes
Jones & Gibbon, 2011	Uganda	Cocoa	Organic	Instrumental variable and multivariate analysis	No
Ruben & Zuniga, 2011	Nicaragua	Coffee	Fairtrade	Propensity score matching, probit regression	Yes
Weber, 2011	Mexico	Coffee	Fairtrade-Organic	Instrumental variable, probit regression	Yes
Blackman & Naranjo, 2012	Costa Rica	Coffee	Organic	Propensity score matching, probit regression	No
Colen et al., 2012	Senegal	Fruit and vegetables	Global GAP	Panel data analysis and OLS	Yes
Jena et al., 2012[9]	Ethiopia	Coffee	Fairtrade-Organic	Propensity score matching and OLS	Yes
Ruben & Fort, 2012	Peru	Coffee	Fairtrade	Propensity score matching, probit regression	Yes
Kleemann & Abdulai, 2013	Ghana	Pineapple	Organic	Propensity score matching, Endogenous switching regression	No

[8] Previous reviews considered here include DeFries et al. 2017, and Oya et al. 2018.
[9] Considered high bias by Oya et al. 2018, but included in DeFries et al. 2017.

Rueda & Lambin, 2013	Colombia	Coffee	Rainforest Alliance	Pair-matched case-control	Yes
Takahashi & Todo, 2013[10]	Ethiopia	Coffee	Rainforest Alliance	Propensity score matching, difference-in-difference panel	Yes
Schoonhoven-Speijer & Ruben, 2014	Kenya	Coffee	UTZ	Multiple regression, logistic regression to control for between-group differences	Yes
Takahashi & Todo, 2014	Ethiopia	Coffee	Rainforest Alliance	Probit model/PSM	No
Chiputwa, Spielman, and Qaim 2015[11]	Uganda	Coffee	Fairtrade, UTZ, Organic	Propensity score matching	Yes
Elbers et al., 2014	Uganda	Coffee	UTZ	Difference-in-difference	Yes
Jena et al., 2015	Nicaragua	Coffee	Fairtrade, Organic	Propensity score matching, endogenous switching regression	No
Rueda et al., 2015	Colombia	Coffee	Rainforest Alliance	Remote sensing analysis, Pair matched case-control	Yes
Cattau et al., 2016	Indonesia	Palm oil	Roundtable on Sustainable Palm Oil	Propensity score matching, analysis of MODIS data	No
Caudill & Rice, 2016[12]	Mexico	Coffee	Bird Friendly	Poisson regression	No
Chiputwa & Qaim, 2016[13]	Uganda	Coffee	Fairtrade, UTZ, Organic	Instrumental variable approach + regression	Yes
Ibanez & Blackman, 2016	Colombia	Coffee	Organic	Matched difference-in-difference	No

[10] Same study region as Takahashi and Todo 2014 and 2017
[11] Same study regions as Chiputwa & Qaim 2016, Meemken et al. 2017
[12] Included in high-level review, but not in the indicator level analysis due to small sample size.
[13]Results for all three certifications are grouped; thus, this is treated as one case when analyzing statistical results.

Karki et al., 2016	India	Coffee	Fairtrade	Panel data analysis, endogenous switching and quantile regression	No
Qiao et al., 2016	China, Sri Lanka	Tea	Fairtrade-Organic	Propensity score matching	Yes
van Rijsbergen et al., 2016	Kenya	Coffee	Far Trade, UTZ	Matched difference-in-difference	Yes
Zulfiqar & Thapa, 2016	Pakistan	Cotton	Better Cotton Initiative	Propensity score matching and probit regression	No
Haggar et al., 2017	Nicaragua	Coffee	UTZ, Rainforest Alliance, Fairtrade-Organic, Fairtrade, C.A.F.E.	Propensity score matching, multiple regression	No
Jena and Grote 2017	India	Coffee	Fairtrade	Propensity score matching	No
Meemken et al., 2017	Uganda	Coffee	Fairtrade-UTZ	Instrumental variable + cross-sectional model	No
Mitiku et al., 2017	Ethiopia	Coffee	Fairtrade, Organic, Rainforest Alliance, Fairtrade-Organic	Propensity score matching + regression	No
Takahashi & Todo, 2017	Ethiopia	Coffee	Rainforest Alliance	Propensity score matching	No
Akoyi & Maertens, 2018	Uganda	Coffee	Fairtrade-Organic, UTZ-Rainforest Alliance-4C	Instrumental variable + regression	No
Carlson et al., 2018	Indonesia	Palm oil	Roundtable on Sustainable Palm Oil	Propensity score matching, panel models	No
Doanh et al., 2018	Vietnam	Tea	Organic	Propensity score matching + regression	No
Froehlich et al., 2018	Brazil	Various	Organic	Propensity score matching, bounded treatment effect, regression	No

Ingram et al., 2018	Ghana, Ivory Coast	Cocoa	UTZ	Propensity score matching and difference-in-difference, supplemented by focus groups and interviews	No
Meemken & Qaim, 2018	Uganda	Coffee	Fairtrade-UTZ	Entropy balancing + regression	No
Minten et al., 2018	Ethiopia	Coffee	Fairtrade, Organic	Propensity score matching, probit model	No
Mitiku et al., 2018	Ethiopia	Coffee	Rainforest Alliance	OLS with control variables, panel fixed effect models	No
Morgans et al., 2018	Indonesia	Palm oil	Roundtable on Sustainable Palm Oil	Propensity score matching and before and after control impact (BACI) analysis	No
Vanderhaegen et al., 2018[14]	Uganda	Coffee	Fairtrade-Organic, UTZ-Rainforest Alliance-4C	Instrumental variable + regression	No
Filho et al., 2019	Brazil	Strawberry	Organic	Propensity score matching, endogenous switching regression	No
Tran & Goto, 2019	Vietnam	Tea	UTZ	Propensity score matching, probit model	No

[14] Same study region as Akoyi & Maertens 2018

Table 2.3: Example of indicator coding.

Source	Pillar	Theme	Country	Certification	Crop	Indicator	Significant difference?	Positive or negative?
(Takahashi & Todo, 2017)	Environmental	Forest quality	Ethiopia	Rainforest Alliance	Coffee	Forest density classification based on NDVI and survey validation[15]	Yes	Positive
(Takahashi & Todo, 2017)	Environmental	Forest quality	Ethiopia	Rainforest Alliance	Coffee	Forest classification in buffer areas (as above, but assesses buffer around certified areas, evaluating spillover effects)	Yes	Positive
Zulfiqar & Thapa, 2016	Economic	Income	Pakistan	Better Cotton Initiative	Cotton	Income	Yes	Positive

[15] Relevant characteristics for each density classification are provided, including number of trees and tree species, height ranges, number of strata, and canopy cover.

2.3. Results

2.3.1 Where are certifications being studied, and how does that compare to the extent of certification globally?

To answer the first research question, attributes of the studies are compared with information on global certification extent from Willer et al. 2019 and Tayleur et al. (2017 and 2018). After determining the proportional coverage of studies for key certifications and crops, several key findings emerge. First, there's a mismatch between what is certified and what is studied. As shown in Tables 2.4 and 2.5, certain crops and certifications are under-represented in the literature - cotton, sugar, cocoa, soy, and palm oil are notable examples, as well as Organic certification. Other crops and certifications are over-represented in the literature, including coffee (a point made in previous reviews), as well as Rainforest Alliance, Fairtrade, and UTZ certification. It should be noted that studies on multiple certification are not reflected in Table 2.5. However, studies of multiple certification occur for the most commonly evaluated standards (e.g. Fairtrade, UTZ, and Rainforest Alliance), so if included would only magnify current trends.

Table 2.4: Comparison of major crops produced under certification[16] with studies of certifications' impacts, sorted by low to high level of representation. A positive difference indicates the crop is under-represented in the literature. Data sources: Willer et al. 2019, Tayleur et al. 2017 (certified crop area estimates), Author's elaboration.

Major certified crops	**Estimated hectares certified (minimum)**	**Percent of total certified area**	**Percent of the crop that is certified**	**Change 2016/2017**	**Change 2013-2017**	**Percent of study coverage**	**Difference between certified area and study coverage**
Cotton	5,154,933	22.20%	16.20%	66.80%	172.40%	1.49%	20.71%
Sugar	1,979,979.00	8.53%	7.60%	88.50%	80.20%	0.00%	8.53%
Cocoa	2,908,640	12.53%	24.80%	22.80%	114.70%	4.48%	8.05%
Soybeans	1,801,269	7.76%	1.50%	-30.20%	-5.90%	0.00%	7.76%
Palm oil	2,537,424	10.93%	11.90%	1.40%	26.10%	4.48%	6.45%
Wheat	1,108,492	4.77%	0.51%			0.00%	4.77%
Other oilseeds	1,002,300	4.32%	0.81%			0.00%	4.32%
Other cereals	549,414	2.37%	0.63%			0.00%	2.37%
Barley	354,881	1.53%	0.72%			0.00%	1.53%
Oats	344,990	1.49%	3.60%			0.00%	1.49%
Pulses	317,446	1.37%	0.40%			0.00%	1.37%
Maize	229,919	0.99%	0.13%			0.00%	0.99%
Fruit & Veg	1,240,463	5.34%	1.03%			4.48%	0.86%

[16] There are known challenges in estimating the total certified area at crop and country level, for example due to multiple certification. See Willer et al. 2019 for more information.

Rice	75,898	0.33%	0.05%			0.00%	0.33%
Root crops	57,928	0.25%	0.10%			0.00%	0.25%
Other	15,433	0.07%	0.17%			0.00%	0.07%
Bananas	340,196	1.46%	6.00%	17.20%	28.60%	2.99%	-1.52%
Tea	668,768	2.88%	16.40%	22.70%	77.30%	5.97%	-3.09%
Coffee	2,533,211	10.91%	23.40%	-8.50%	8.70%	74.63%	-63.72%

Interestingly, the top five understudied certified crops - cotton, sugar, cocoa, soybeans, and palm oil - are all known to drive significant and urgent sustainability issues. These range from severe labor risks, intensive pesticide and water use, deforestation, and land conversion to additional human rights and farmer livelihoods issues. Research on these crops appears to be increasing, though, as all the identified studies which evaluate them were published in 2016 or later. As shown in Table 2.4, cotton has the most significant discrepancy between its estimated certified area (over 5 million hectares, about 22% of total certified area) and study coverage. Only one study was identified assessing the impact of the Better Cotton Initiative, the predominant cotton certification standard (Zulfiqar & Thapa, 2016). Based on this review, it is possible that two of the most underrepresented certified crops - sugar cane and soybeans - have never been assessed in the peer reviewed literature using a robust impact evaluation method. In addition, cotton, sugar cane, and cocoa are all experiencing significant increases in certification, heightening the need for robust impact evaluation. Certified cotton area increased by over 170% from 2013-2017, sugar certified area increased by 80%, and cocoa certified area increased by 115% (Willer et al. 2019). From a standards perspective, Organic faces the largest discrepancy between certified area and study coverage. IFOAM organic cropland covered approximately 69 million hectares in 2017, representing 72% of certified area. It is particularly important to evaluate trade-offs when considering Organic certification, since it has been found to produce benefits such as improved soil health and biodiversity, while also negatively impacting productivity and income (Vanderhaegen et al., 2018). Most of the other certifications with the largest gaps in evaluation studies - Proterra, Better Cotton Initiative, Global GAP, Cotton Made in Africa, Roundtable on Responsible Soy, and Bonsucro - conduct their own performance monitoring to measure progress on certified farms, but do not use a counterfactual approach to assess their impact over time.

Table 2.5: Comparison of major certifications globally with studies of certifications' impacts, sorted by low to high level of representation[17]. Data sources: Willer et al. 2019, Roundtable on Sustainable Biomaterials 2018, Author's elaboration.

Area certified for 13 major standards globally	**Hectares certified (2017)**	**Percent of total certified area**	**Percent of study coverage**	**Difference between certified area and study coverage**
Organic cropland (IFOAM)	69,845,243	71.50%	20.90%	50.56%
Proterra	2,339,259	2.40%	0%	2.39%
Better Cotton Initiative	3,561,000	3.60%	1.50%	2.15%
Global GAP	3,548,194	3.60%	1.49%	2.14%
Cotton Made in Africa	1,619,469	1.70%	0%	1.66%
Roundtable on Responsible Soy	1,259,672	1.30%	0%	1.29%
Bonsucro	1,161,000	1.20%	0%	1.19%
Roundtable on Sustainable Biomaterials	18,100	0.01%	0%	0.01%
Roundtable on Sustainable Palm Oil	3,301,088	3.40%	4.50%	-1.10%
Common Code for the Coffee Community (4C)	1,630,546	1.70%	3.00%	-1.32%
Rainforest Alliance/Sustainable Agriculture Network	3,458,167	3.60%	11.90%	-8.40%
UTZ	3,376,870	3.50%	13.40%	-9.98%
Fairtrade	2,634,678	2.70%	26.90%	-24.17%

[17] One study looked at Bird Friendly certification, not shown here (<15,000 ha certified). The C.A.F.E. practices standard is also not shown here.

The analysis also reveals important insights about regional and country-level research coverage and gaps. As illustrated by Table 2.6, Africa is the most common region of analysis, followed by Latin America (51% and 34% respectively), with Southeast Asia, South Asia, and East Asia representing about 7%, 6%, and 1% of studies. No studies were identified assessing certification in North America or Australia, although certification does take place in these areas. There are 19 countries covered by the included studies - the most common countries are Uganda with 24% of cases, Ethiopia with 16% of cases, and Nicaragua with 13% of cases. Peru, Costa Rica, Colombia, Kenya, and Indonesia each represent about 4% of the reviewed cases. Table 2.6 presents the top 13 countries that appear to be under-represented in the literature. This list represents a mix of high, middle, and low-income countries on six continents, and includes Brazil, Australia, Malaysia, Indonesia, the United States of America, Canada, Zambia, and the Ivory Coast. There are also several European countries currently under-represented in the literature (Spain, Italy, France, and Germany). These countries likely have significantly different certification compositions. For example, the United States' certified area estimate is largely driven by barley, while other countries like Australia, Brazil and Indonesia have a larger mix of crops. The full country comparison list can be found in the Supplementary Materials. Comparing this list to country-level certified area based on Tayleur et al.'s 2017 analysis of certification extent helps illuminate current research needs. Figure 2.3 depicts this information visually on a map at country level. Comparison to Tayleur et al.'s 2018 map of certification extent allows identification of clustered certified regions which appear to have not yet been evaluated. While this does not reflect all certified area, it appears that large areas of certified coffee, sugar, and soy in Brazil, cocoa produced in the Ivory Coast, and oil palm produced in Malaysia have not yet been evaluated in the peer-reviewed literature. Figure 2.4 highlights the estimated certified areas for cocoa and coffee in the Ivory Coast and Ghana, as well as several certified crops in Brazil, using 30km x 30km grid cells from Tayleur et al. 2018.

Table 2.6: Abridged comparison of countries with certified production globally with studies of certifications' impacts, sorted by low to high level of representation. Data sources: Tayleur et al., 2017 (certified area estimates), World Bank 2020, author's elaboration.

Country	Certified area (ha, Tayleur et al. 2017)	Percent of total cert. area	Percent study coverage	Difference between certified area and study coverage	Income level
Brazil	2,386,045.00	15.93%	3%	12.94%	Upper middle-income
Australia	653,733.50	4.36%	0%	4.36%	High income
Malaysia	649,866.00	4.34%	0%	4.34%	Upper middle-income
Ivory Coast	861,866.40	5.75%	1%	4.26%	Lower middle-income
Spain	596,629.00	3.98%	0%	3.98%	High income
Italy	562,749.60	3.76%	0%	3.76%	High income
United States of America	546,591.50	3.65%	0%	3.65%	High income
Zambia	318,680.00	2.13%	0%	2.13%	Lower middle-income
France	309,019.50	2.06%	0%	2.06%	High income
Turkey	307,157.10	2.05%	0%	2.05%	Upper middle-income
Canada	296,972.60	1.98%	0%	1.98%	High income
Germany	268,272.20	1.79%	0%	1.79%	High income
Indonesia	931,536.90	6.22%	4%	1.74%	Lower middle-income

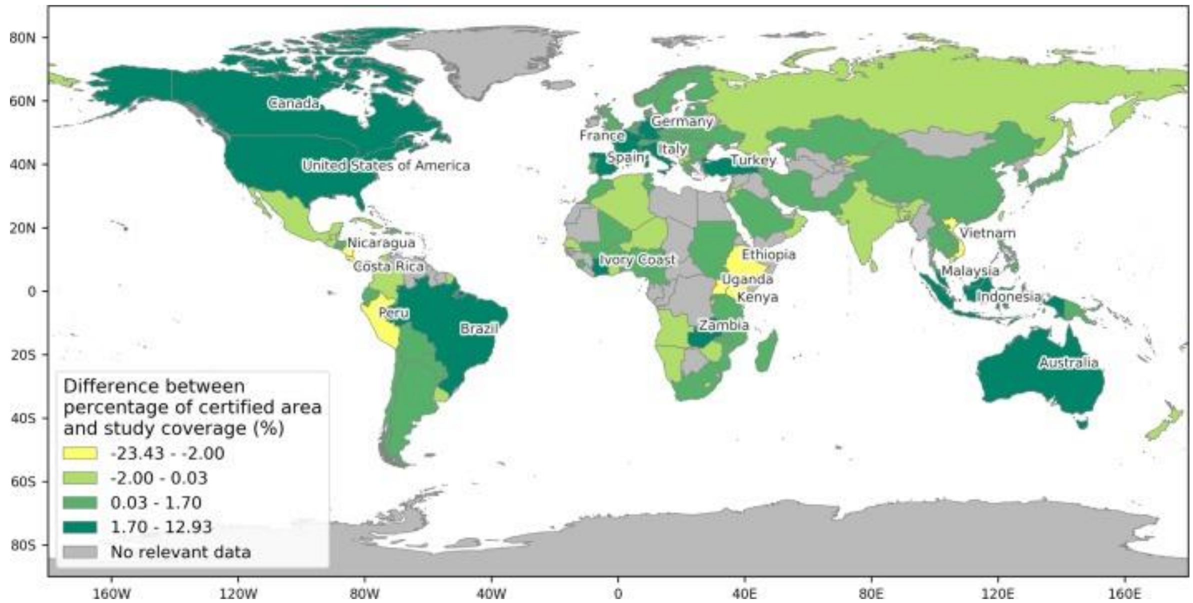
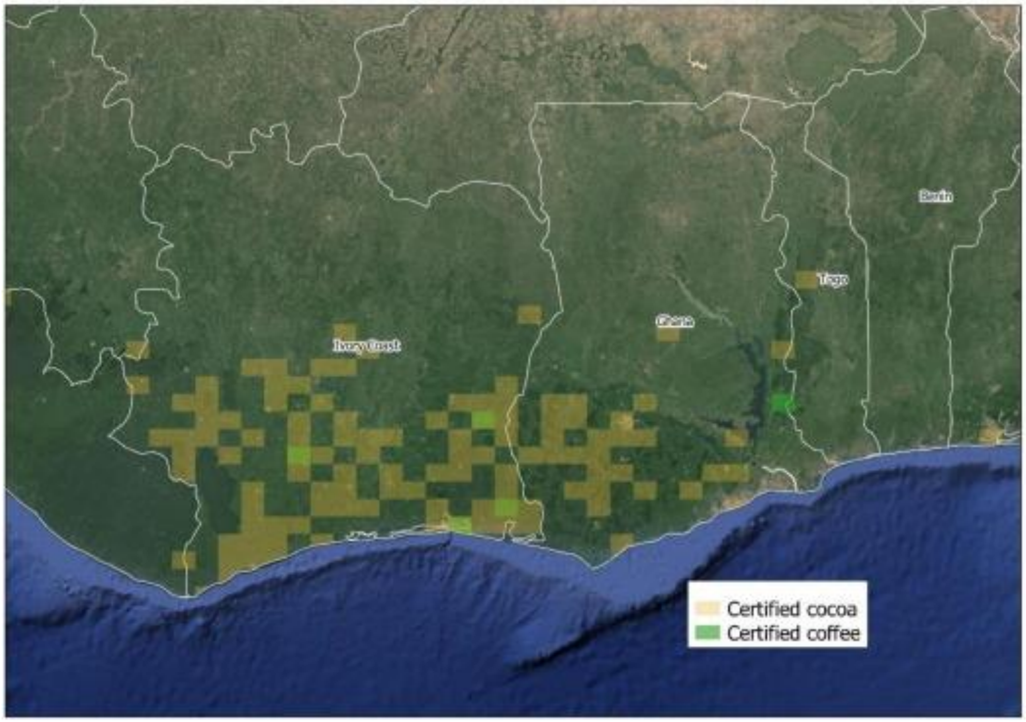


Figure 2.3: Country research gaps in evaluations of certified agriculture. Higher values indicate under-represented countries, with the yellow and darkest green classes representing the areas of highest and lowest representation, respectively. Countries falling into these two classes are labeled. Data sources: Tayleur et al. 2017 and author's elaboration.



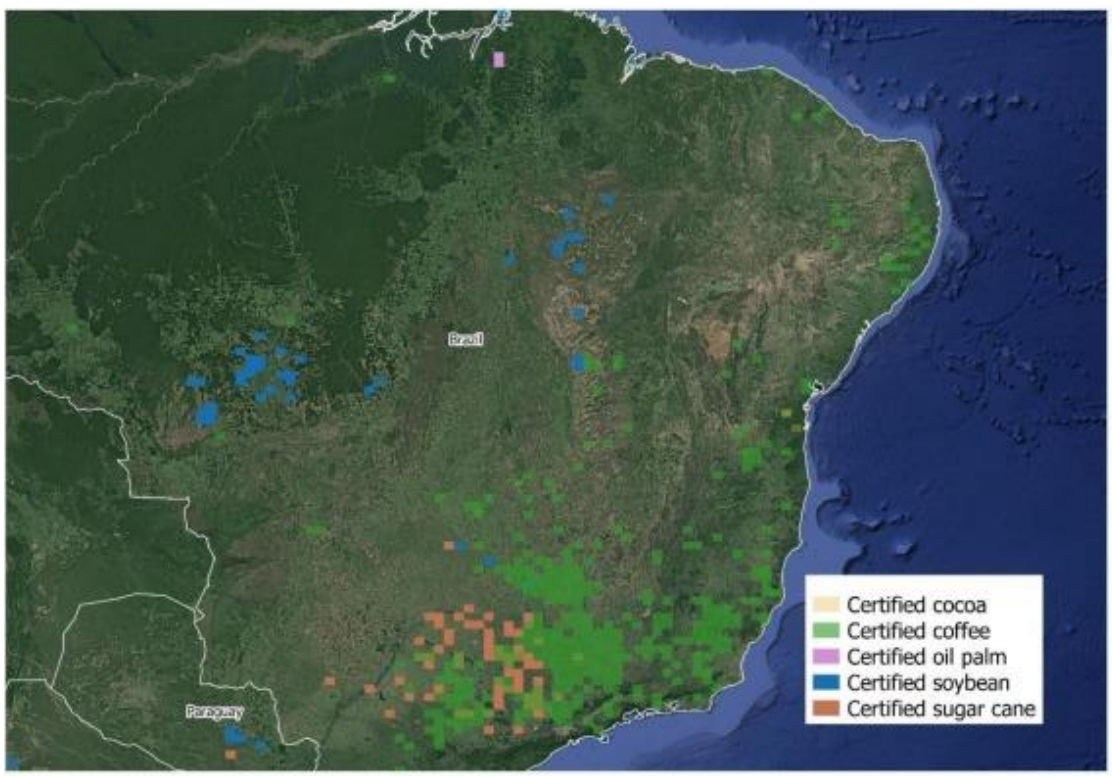


Figure 2.4: Spotlight on under-studied countries producing certified agriculture. Certified area for major crops (Tayleur et al. 2018) are displayed for the Ivory Coast and Brazil. Data sources: Tayleur et al 2017b and author's elaboration. **NB:** There are known challenges with identifying precise locations of certified area.

2.3.2 Which pillars of sustainability are included in evaluation studies, and which indicators are used to measure success?

The economic pillar is studied the most frequently (84% of the reviewed articles). There is less frequent analysis of social outcomes (43% of articles), as well as environmental outcomes (43% of articles). Only 20% of the studies (nine in total) consider all three pillars, while 30% look at two pillars (typically, economic and social outcomes together).

The analysis provides some important insights regarding environmental indicators. First, the majority of indicators measure "outputs", or direct results from the certification (60% of the reviewed environmental indicators). These are often practice-focused, for example related to agrochemical use, and best management measures related to crops, soil, trees, or water resources. A significant portion of the indicators, about 40%, measure outcomes, or what could be considered "ultimate goals" of certification related to environmental conditions. Outcome indicators focus on topics including forests and trees (e.g. density, diversity, tree cover, forest quality), biodiversity (invertebrate abundance), carbon stocks (tree biomass, soil organic carbon, and total carbon stocks), deforestation, and fire activity (for palm oil). Notably, no direct measurements were employed for water resources, and only a few outcome indicators for soil incorporated direct measurement. For biodiversity, in addition to invertebrate diversity (Vanderhaegen et al., 2018), one study analyzes orangutan presence (Morgans et al., 2018).

Additionally, the majority of environmental indicators (58%) are measured using survey data, 33% are measured using field observations, and 8% rely on remotely sensed data (e.g. LANDSAT and MODIS). In terms of the themes covered by environmental indicators, the most common are best management practices (23% of environmental indicators) and agrochemical and input use (20% of indicators). Tree density and diversity is the third-most common theme, followed by carbon storage, soil, and biodiversity (9%, 7%, and 7% respectively). Water conservation is only considered in 2% of the indicators (and only for output indicators), and habitat quality for 3% of the indicators. For additional information on the environmental indicator coding, see the Supplementary Materials.

For social sustainability, about 56% of indicators measure outputs or direct results of the certification, and about 44% measure ultimate outcomes. Some of the most common themes covered by social indicators include 1. perception and satisfaction with 14% of the indicators (e.g. regarding technical assistance, cooperatives, and economic well-being); 2. social networks (12% of indicators), covering topics like participation in farming organizations; 3. nutrition and food security, including the extent of nutrient, vitamin and energy deficiencies (11% of indicators); and 4. gender, e.g. control of resources, assets and decision-making within the household, and participation in agriculture (10% of indicators). Additional social themes include poverty, labor, health, education, and child labor.

The reviewed studies provide useful examples of outcome measurement for themes like education, health, nutrition and food security, and gender. For example, indicators to measure educational outcomes include school attendance and the maximum grade obtained for school-age children, as well as educational expenditures (Arnould et al., 2009, Minten et al., 2018, Meemken et al., 2017). Arnould et al., 2009 use two health indexes to analyze health outcomes - one which measures the extent to which individuals in the household suffered specific illnesses or death, and one measuring receipt of treatment for those illnesses. Chiputwa & Qaim, 2016 and Meemken et al., 2017 both assess nutritional outcomes, primarily by measuring per adult equivalent consumption of calories and micronutrients, using these values to estimate the prevalence and depth of any deficiencies.

In terms of the economic indicators, the measurement of outputs and outcomes is fairly equal, with 48% and 52% of the indicators, respectively. Income is the most common economic sustainability theme, representing 25% of the indicators; it is followed by productivity (12% of indicators), price (9% of indicators), and expenditures (7% of indicators). Income indicators often include overall household income as well as income from the certified crop, and yield is typically measured in volumes per hectare or per tree (for coffee). Several studies also go beyond income to measure the prevalence of poverty among certified and non-certified groups, which can help illustrate systemic challenges to well-being, despite potential improvements in economic output indicators like price and volumes sold (e.g. Mitiku et al., 2017, Jena and Grote 2017, Akoyi & Maertens, 2018, Vanderhaegen et al., 2018).

2.3.3 What does the current evidence base suggest regarding outcomes of voluntary agricultural sustainability standards?

What conclusions can we draw from this evidence base regarding outcomes of voluntary agricultural sustainability standards? When grouped by case, the most common results are positive (51%), no difference (41%), and negative (8%). There are no significant differences in terms of the average proportion of positive and negative results when grouped by sustainability pillar. The proportion of no difference results is significantly different between economic and social pillars as measured by the Kruskal-Wallis test, with the social pillar tending to exhibit a higher proportion of no difference results (see Table 2.7 and Figure 2.5). Kruskal-Wallis tests also suggest no significant differences between a selection of sustainability themes, although negative results border on significantly different (see Table 2.8 and Figure 2.6). Sustainability themes with the highest proportion of positive results include environmental outcomes, income, agrochemical and input use, and social networks; conversely, best management practices and perception and satisfaction show high proportions of not significant results, and gender and productivity show the highest amounts of negative results. Negative results for productivity occur primarily for Organic and Fair Trade-Organic certification. While it is possible to further analyze indicator results across certifications and crops, this is likely only advisable when there have been a sufficient amount of studies conducted on the certification or crop in question, to reduce the likelihood of erroneous conclusions.

It is important to treat these results with caution due to the still-limited evidence base. This is particularly true for two reasons - first, that evaluation studies have been largely unanimous in their emphasis on context in driving or enabling certification outcomes, and second, that 75% of the reviewed studies focus on coffee. In addition, participants may perceive the benefits they receive from certification differently than outcome indicators suggest, which can be captured through qualitative methods and direct feedback from participants. One example of this is found in Jena et al.'s 2015 study of Fairtrade certification in Nicaragua, which found no difference between certified and non-certified farms in terms of overall income, although participants reported benefits related to education and health services not reflected in outcome indicators.

Table 2.7: The average proportion of positive, negative and no difference results when comparing between certified and uncertified participants, grouped by case for each pillar of sustainability. Standard deviations are shown in parentheses. Differences between sustainability pillars were not significant, apart from the proportion of non-significant results (p=0.017 for Kruskal-Wallis test). The number of indicators per case varies between 1-57; the average is 11.

Pillars of Sustainability	**Number of cases**	**Proportion positive**	**Proportion negative**	**Proportion no difference**
Environmental	26	0.47 (0.36)	0.06 (0.21)	0.47 (0.35)
Social	29	0.33 (0.36)	0.07 (0.16)	0.6 (0.36)
Economic	55	0.51 (0.38)	0.12 (0.27)	0.38 (0.33)
All	62	0.51	0.08	0.41

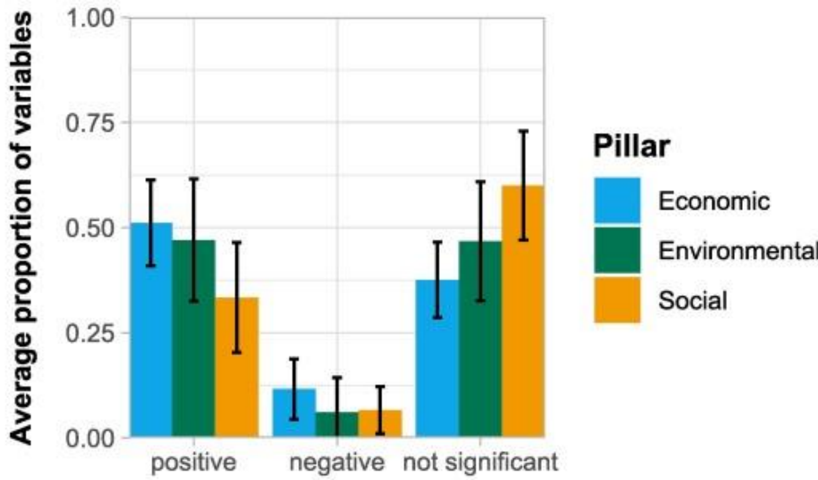


Figure 2.5: The average proportion of positive, negative and no difference results for comparison between certified and uncertified participants, when grouped by case, for each pillar of sustainability. Error bars show 95% confidence intervals.

Contextual factors that have been identified as important influencers of certification's outcomes include: the prevalent poverty and livelihoods conditions (e.g. severity of poverty and dependence on farm income prior to the intervention); market structure (the existence of a price premium and cooperative structure, for example, as well as the contract type); market conditions (e.g. sales to certified markets and global commodity price trends); and the extent to which certification requirements surpass legal requirements (Jena et al., 2015, Minten et al., 2018, Ruben et al., 2009, Qiao et al., 2018, Oya et al., 2018, van Rijsbergen et al., 2016). Certifications operate within in a broader market environment, and there is a risk that severe shocks caused by price volatility can mask the benefits of years of productivity gains, depending on the support systems available to producers (Dave McLaughlin, personal communication). On the environmental side, contextual conditions could include broader environmental trends, for example the regional deforestation level (Rueda et al., 2015, Carlson et al., 2018). Different standards also vary in their requirements, as well as the training, assistance, and other benefits they provide to producers - this is important to keep in mind when attempting to generalize results. For example, Fairtrade certification includes social premium funds that go to the community, to improve services like education, health care, and infrastructure (Karki et al., 2016). Standards also vary in the extent to which they address environmental issues like reduction of greenhouse gas emissions and deforestation (Tayleur et al. 2017, Table 2.1).

Table 2.8: The average proportion of positive, negative and no difference results between certified and uncertified participants, grouped by case for a selection of indicator themes. Standard deviations are shown in parentheses. Differences between sustainability themes were not significant as measured by the Kruskal-Wallis test, although negative results bordered on significantly different (p=0.052).

Indicator theme	**Number of cases**	**Proportion positive**	**Proportion negative**	**Proportion no difference**
Income	32	0.54 (0.47)	0.08 (0.26)	0.38 (0.44)
Productivity	29	0.37 (0.44)	0.24 (0.41)	0.39 (0.45)
Environmental output	19	0.37 (0.38)	0.09 (0.24)	0.54 (0.40)
Environmental outcome	16	0.63 (0.35)	0.02 (0.08)	0.35 (0.32)
Best management practices	15	0.26 (0.40)	0.08 (0.15)	0.66 (0.40)
Poverty	13	0.38 (0.51)	0.00	0.62 (0.51)
Social network	9	0.41 (0.41)	0.15 (0.34)	0.44 (0.42)
Agrochemical and input use	8	0.50 (0.37)	0.19 (0.37)	0.31 (0.34)
Perception and satisfaction	8	0.16 (0.27)	0.18 (0.35)	0.66 (0.39)
Gender	7	0.18 (0.41)	0.39 (0.45)	0.43 (0.46)

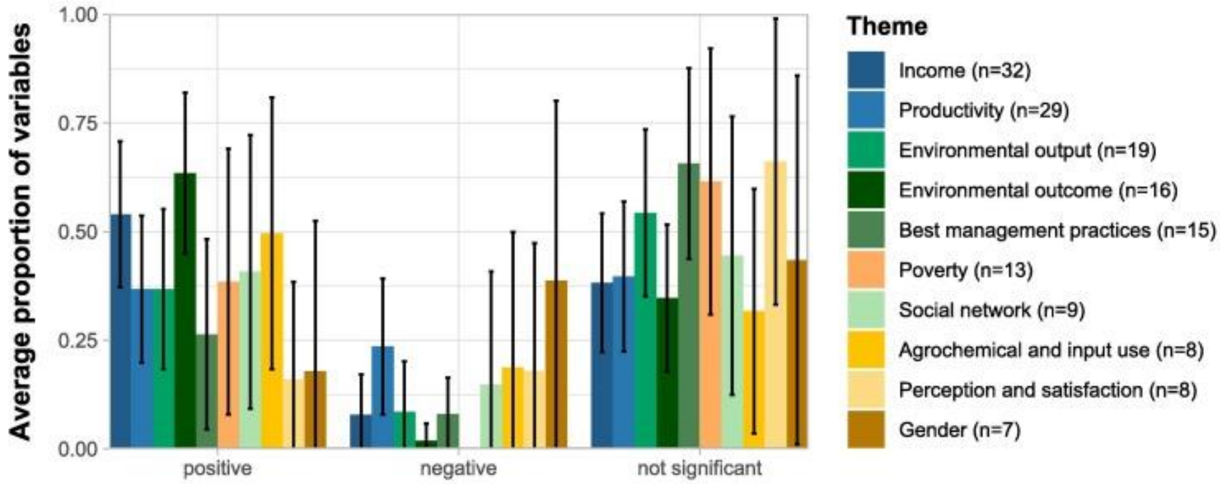


Figure 2.6: The average proportion of positive, negative and no difference results between certified and uncertified participants, when grouped by case, for a selection of sustainability themes. Error bars show 95% confidence intervals.

2.3.4 A qualitative assessment of best practices and novel examples

In addition to addressing the key research questions, this analysis helps identify best practices in assessing certification's impacts. First, there are several studies which provide strong examples of multi-pillar evaluations. The findings indicate that only 20% of studies look at all three pillars of sustainability; of these, five were published prior to 2015, and four were published in 2018 (Fort & Ruben, 2009; Ingram et al., 2018; Minten et al., 2018; Morgans et al., 2018; Ruben et al., 2009; Ruben & Fort, 2012; Ruben and Zuniga, 2011; Schoonhoven-Speijer and Ruben, 2014; Vanderhaegen et al., 2018). Some specific strengths of multi-pillar studies include consideration of trade-offs between environmental and socioeconomic outcomes, as well as data collection regarding farmer's perceived benefits of certification. Another important area of progress has been the consideration of nuanced social indicators, including issues like health, education, gender, and nutrition (see Chiputwa & Qaim, 2016, Meemken & Qaim, 2018). These studies incorporate data collection approaches from the development field (among others), enabling assessment of intra-household dynamics as well as human well-being outcomes.

Finally, there have been significant developments in the assessment of environmental sustainability indicators, particularly using field measurements and remotely sensed imagery (e.g. Haggar et al., 2017, Vanderhaegen et al., 2018, Takahashi and Todo (2013, 2014, and 2017) and Cattau et al., 2016). These studies employ clear sampling strategies and analytical approaches to address various sources of bias, and data collection protocols to enable comprehensive and efficient field measurements. Together, these best practices enable authors to identify novel findings related to the outcomes of certification, and it merits consideration why more studies do not replicate these approaches. These best practices are elaborated upon further in the Discussion section.

2.3.5 Limitations

Ideally, a review of this type would include a step of cross-checking methodological decisions, including the process for searching for articles, selecting articles for inclusion, and coding indicator results, to ensure inter-rater reliability (Siddaway et al., 2019). However, this is not always possible, as was the case for this review. Here, this is addressed by relying closely upon previous reviews, and providing detailed information on methods. Methodological details are provided in the Supplementary Materials to enable reproducibility, and it assumed that any small divergences in article selection or coding would not greatly affect key trends in the results.

Additional limitations of this review include that it was conducted in the English language, and that it did not consider evaluations which had not been peer reviewed. There is a large amount of grey literature focused on the impacts of VSSs, and some of this literature likely applies a counterfactual approach. Future reviews could aim to broaden the search to additional languages beyond English and consider rigorous grey literature sources in addition to peer-reviewed articles.

2.4 Discussion

2.4.1 Relating findings to other recent reviews

Among the papers that have robustly evaluated the impact of certification, there have been mixed findings regarding outcomes. DeFries et al. 2017 found that the majority (58%) of outcome indicators in evaluated studies found no significant difference between certified and non-certified producers, while 34% of indicators represented positive outcomes (DeFries et al., 2017). When grouped by case, this study finds a slightly more positive trend, the average proportion of positive indicators was .51. Although economic and environmental pillars have a higher average proportion of positive results than social indicators, these differences were not statistically significant. However, the fact that social indicators tended to exhibit no difference between certified and non-certified producers may suggest that sustainability standards are less effective in driving change for social sustainability concerns (see Ingram et al. 2018).

Oya et al. (2018) found that most studies which looked at farm income effects found positive and statistically significant results, although there was variation in individual study-level effects. The authors surmised that this was driven by specific capacity-building activities, farm productivity, and market conditions (e.g. prices received), with the price linkage as key to overall effects on income (Oya et al., 2018). There is also a challenge translating farm income effects into broader household income effects, due to contextual factors, and household dependence on other income sources (Oya et al., 2018, DeFries et al. 2017) – here, we see this reflected in the lower proportion of positive results for poverty indicators as compared to income indicators. A more recent review (Meemken 2020) also found evidence of positive price and income effects. Overall, this study finds that income indicators tend to be positive; the only theme with a higher average proportion of positive results was the "environmental outcome" category.

2.4.2 Recent progress - multi-pillar studies, increased consideration of trade-offs, and non-economic benefits to producers

Despite persistent research gaps, significant progress has been made in the evaluated studies, which can present further opportunities for learning and replication. Some of the key areas of progress have been an increase in the number of multi-pillar studies, and the inclusion of important socioeconomic variables beyond price and income. The discussion below highlights a few exceptional studies which provide examples of best practices.

First, in their analysis of Fairtrade and Organic certification in Ethiopia, Minten et al., 2018 explore the receipt of price premiums by certified producers, as well as environmental and social outcomes. In addition to the price premium, outcome indicators include agricultural management practices (e.g. stumping, compost use, and use of agrochemical inputs), as well as school attendance for school-aged children (Minten et al., 2018). The authors also ask farmers about the benefits they see in engaging with cooperatives more broadly. Although this study finds limited evidence of income improvement due to certification, the inclusion of social benefits like school attendance enables the recognition of non-economic benefits that may influence farmers' decisions to participate in certification, in addition to their households' overall well-being (Minten et al., 2018).

Vanderhaegen et al., 2018 provide another useful example of a multi-pillar analysis, in their study of multiple coffee certifications (Fairtrade-Organic and UTZ-Rainforest Alliance-4C) in Uganda. The indicators included in this study are unique among those reviewed, as they span socio-economic issues (e.g. poverty, income, and yield) as well as practice and outcome-based environmental issues. Within the environmental pillar, this study addresses several important themes including biodiversity, carbon storage, and tree density and diversity, with data collected through field observations, in addition to the agricultural best management practices that are typically included in survey-based approaches to evaluate certification (Vanderhaegen et al., 2018). The authors go one step further and conduct a correlation analysis to illuminate trade-offs between sustainability pillars. This approach underpins a key finding - the two multiple certifications have different implications for the level of coffee intensification, with UTZ-RA-4C increasing the likelihood of using agrochemicals (along with practices like use of shade trees and intercropping with legumes), and Fairtrade-Organic reducing the use of inputs and increasing likelihood of soil tillage (Vanderhaegen et al., 2018). The two schemes also have different outcomes for socio-economic performance at farm level and prices, as well as environmental variables like carbon stocks and tree cover density (Vanderhaegen et al., 2018). The findings emphasize that increased agrochemical inputs which tend to occur in UTZ-RA-4C certification lead to yield, labor productivity and income increases, but decreases in invertebrate abundance and diversity (Vanderhaegen et al., 2018). This critical finding regarding trade-offs between yields and ecosystem services would not have been possible with a single pillar approach, which further underlines the importance of applying a multi-pillar methodology to other crops, certifications, and production contexts. This is also important because some studies have suggested that productivity increases are more important than price premiums for increasing overall returns to farmers (Akoyi & Maertens 2018).

Another important area of progress has been the consideration of nuanced social indicators, including issues like health, education, gender, and nutrition (see Chiputwa & Qaim, 2016, Meemken & Qaim, 2018). In their analysis of Fairtrade-UTZ certification in Uganda, Meemken & Qaim (2018) include indicators like asset ownership for female household members, as well as a 24-hour time recall measuring division of labor, a subjective measure regarding satisfaction with time for leisure activities, and interactions with farmer organizations, extension officers, and training sessions to capture rural services. This approach enables the authors to identify an increase in total household assets for certified female-headed households as compared to non-certified, due to higher coffee revenues. The authors also found evidence that standards impacted the distribution of assets within households, with a positive effect on joint asset ownership within male-headed households (Meemken & Qaim, 2018). Further, indicators related to social networks – for example, organizational participation and measurement, including farmers' organizations – also suggest some positive benefits for certified producers (see van Rijsbergen et al. 2016, Meemken & Qaim, 2018). This lends support to growing evidence that price premiums are just one piece of the puzzle in terms of certification's socioeconomic impacts.

While there have been significant improvements in the exploration of VSS impacts for multiple pillars of sustainability, studies which analyze all three pillars are still the minority in the literature. Conducting evaluations which incorporate social, economic, and environmental components should continue to be a priority.

2.4.3 Acting on research gaps - representation of different standards, crops, and regions; environmental outcome measurement

The key research gap indicated by this review is an under-representation of studies for certain crops, certifications, and regions. This is particularly important since we do not have any reason to believe that findings for coffee - which represents 75% of the cases included - will hold true for other crops. Additionally, several of the main understudied crops - cotton, cocoa, sugar cane, soy, and palm oil - are known to cause severe and urgent sustainability issues.

Why should the high proportion of studies focused on coffee give us pause in generalizing results? There are several ways in which coffee certification substantively differs from certification for other crops. Out of the 13 major agricultural sustainability standards, seven certify coffee. Coffee is shade-tolerant, meaning that it can prosper under full or partial shade (Rainforest Alliance, 2017). For this reason, it is fairly unique among the major certified crops, and shade-grown coffee can bring significant environmental and socioeconomic benefits (De Beenhouwer et al., 2013, Jha et al., 2014). The share of shade-grown coffee has been decreasing over time, and several key production countries now rely primarily on full-sun coffee production systems (Jha et al., 2014). Still, many of the major certified crops including cotton, sugar, soybeans, and palm oil do not have an equivalent to shade grown coffee - while production systems may represent differing levels of intensity, agroforestry systems are not common at the global scale. This means that we can expect to see different environmental issues and outcomes for those crops. As one example of this, the Rainforest Alliance standard sets a minimum canopy cover requirement of 40% for their certified coffee; the only other crops with canopy cover minimums are cocoa (30%), clove and vanilla (40%), and pepper (20%) (Rainforest Alliance 2017). Interestingly, no studies were identified evaluating certified coffee in Brazil, which has one of the highest certified coffee areas (Willer et al. 2019). Further, coffee faces specific market conditions that distinguish it from other crops. While it represents one of the most well-developed markets for sustainably certified agriculture, it also faces challenges, including a mismatch between supply and demand. A 2014 report found that although 40% of global coffee production complied with global standards, only 10% was sold as such (Potts et al. 2014). Finally, we can expect that the willingness to pay a price premium will differ significantly for coffee as compared to a commodity crop like soy or wheat, which may have implications for financial returns to producers.

An additional gap is a lack of direct measurement for environmental outcomes. While 38% of reviewed studies considered environmental indicators, only 22% explicitly considered environmental outcomes. While information on the use of specific agricultural practices is helpful, it is important to complement this with direct measurement as well as the measurement of true outcomes, rather than focusing exclusively on outputs (Milder et al., 2015). Despite this limitation and the overall lower proportion of studies looking at environmental outcomes, there have been several innovative approaches to the topic. The two main methods for measuring environmental outcomes incorporate field observations and direct measurement, and remotely sensed data.

The use of field measurements for environmental outcomes is emblemized by studies from Haggar et al., 2017 and Vanderhaegen et al., 2018. The first study aligns with the Committee for Sustainability Assessment (COSA) method for assessment of coffee sustainability, utilizing indicators that can be measured by trained evaluators (rather than scientific experts), and can be completed in half a day to one day per farm, thus enabling larger sample sizes. Environmental outcomes assessed by this study include tree density and diversity, habitat quality, and carbon storage (Haggar et al., 2017). This balance of rigor and feasibility should continue to be adopted in future studies, including for under-studied crops like cotton, sugar, cocoa, soybeans, and palm oil. Vanderhaegen et al., 2018 collected environmental data in a subsample of 74 coffee fields, using stratified random sampling based on elevation and soil type. These fields were then matched with similar non-certified fields using propensity score matching, and field measurements were completed for half-hectare plots randomly placed throughout fields (Vanderhaegen et al., 2018). Field measurements included diameter at breast height and height for plant species; woody debris, stem and plant counts; litter collection; and soil samples for bulk density and soil organic carbon (Vanderhaegen et al. 2018). This study is also notable in that it included outcome variables related to soil and biodiversity (invertebrate abundance and diversity) (Vanderhaegen et al. 2018).

Takahashi and Todo (2013, 2014, and 2017) and Cattau et al. (2016) provide useful examples of evaluating environmental outcomes using remotely sensed imagery. Takahashi and Todo (2014) look at Rainforest Alliance certification of coffee in Ethiopia, considering outcomes related to forest conservation, deforestation, and forest quality. The study utilizes two years of Landsat 7 imagery (2005 and 2010) and a probit model, supplemented by a household survey to understand the relationship between forest conservation and socioeconomic characteristics. The authors find that certification has a positive effect on forest conservation; conservation was also affected by years of formal education and the total area of agricultural land at the household level (Takahashi & Todo, 2014). They also find evidence that certification had a significant impact on behavior for producers with more limited assets (Takahashi & Todo, 2014). Cattau et al. (2016) evaluate palm oil concessions in Indonesia to determine whether Roundtable on Sustainable Palm Oil certification impacts fire activity. This analysis uses nonparametric matching methods and data from Global Forest Watch and MODIS Active Fire Detections, finding a positive and significant relationship for one out of four outcome variables (density of fire activity on non-peatlands in wet years), with the remaining variables showing no significant difference between certified and non-certified concessions (Cattau et al., 2016).

2.4.4 Challenges for robust evaluations of certification's impacts

There are clear challenges in designing and implementing robust impact evaluations of certification. One such challenge is the resources needed to implement them (Margoluis et al., 2009). Studies typically rely on field-collected survey data, sometimes in combination with secondary data sources, like government agricultural censuses or remotely sensed imagery. In order to conduct robust statistical analysis, field surveys must cover a sufficient number of farms. Sample design is also complex, with many studies using two or three stage sampling designs, with careful matching between treatment and control groups to account for selection bias. The resource requirements of these studies are significant and the logistics are not uncomplicated, which is characteristic of environmental impact evaluations more broadly (Mascia et al., 2014).

Several additional difficulties of designing robust impact analyses of certification have been identified. These include considerations related to control group selection. For example, studies may include producer communities in close proximity to the intervention group as a control. However, spillover benefits have been found to affect outcomes for nearby communities as well as those receiving certification - this may create an impression of a smaller difference between treatment and control groups, if not identified and appropriately highlighted in results (Komives et al., 2018, Jones & Gibbon, 2011). Some studies have accounted for this by selecting control groups outside of the influence of certification activities, or explicitly considering spillover effects in data collection (see Ingram et al., 2018, Van Rijsbergen et al., 2016). Additionally, differences in timing of certification among certified groups may create issues with identifying trends in the data (Komives et al., 2018).

Another challenge of building the evidence base has been indicator consistency and quality. Milder et al. (2015) emphasized that use of different outcome variables between studies presented a challenge for robust comparative analysis; in addition, studies of environmental changes tended to focus on management practices, rather than environmental outcomes per se (Milder et al., 2015). The authors referenced several programs which sought to define common indicators, including the Committee on Sustainability Assessment (COSA) and the SAI Platform, and advocated for the development of common metrics (Milder et al., 2015, COSA 2020, SAI Platform, 2014). Indicator consistency involves a delicate balance, as production systems between small and large producers vary significantly, and necessitate different measures of success (Dave McLaughlin, 2020, personal communication).

2.5 Conclusions

VSSs represent a huge investment in sustainability from the international community. Given the extent of current sustainability problems, it is more important than ever to understand the outcomes of these interventions for local communities and ecosystems. This review articulates several important priorities for the evaluation of voluntary sustainability standards' impacts moving forward. First, future research should build on recent gains in terms of assessing multiple pillars of sustainability. There is evidence that trade-offs can occur between socioeconomic and environmental outcomes - failing to consider these trade-offs will result in an incomplete picture of the benefits and challenges of certification.

Secondly, continued efforts to standardize and align indicators are critical to inform future reviews and meta-analyses of certification. It is important that outcome indicators reflect actual goals of certification, and not contextual factors that cannot reasonably be influenced by the program. While price premiums are a crucial component of benefits to producers, it is also important to robustly evaluate other socioeconomic impacts of certification and producers' perceptions of these impacts, which may influence producers' decisions to participate and stay in the program.

Third, future research should continue to build on progress to measure environmental outcomes. Although environmental outcomes are not frequently evaluated, there have been significant developments in direct measurement and remote sensing methods. Studies should continue to build on these gains to determine whether certifications are achieving their goals. Remote sensing has particular benefits for assessing landscape-level impacts of certification, although there can be challenges in defining suitable proxies for sustainability outcomes at the landscape level and addressing confounding factors (Rueda et al., 2015, Morgans et al., 2018). This study's findings generally align with DeFries et al. (2017) at the indicator level, but paint a slightly more positive picture, as the average percentage of positive indicators was 51%. However, these trends should be interpreted with caution, and a more balanced evidence base is required to increase confidence in the findings.

Future research should aim to address the research gaps identified within this review, including evaluation of under-represented crops, standards, and geographical areas. Due to the growing acknowledgement that certification alone is not sufficient to achieve sustainability goals, there is also a need to evaluate the adoption and impact of VSSs in combination with other approaches, including landscape-based conservation and policy support for producers. Although growing, the limited evidence base has significant implications for continuous improvement of standards, their potential integration with other policy or program interventions, and their ability to deliver on key sustainability objectives.

2.6 Bibliography

Akoyi, K. T., & Maertens, M. (2018). Walk the Talk: Private Sustainability Standards in the Ugandan Coffee Sector. The Journal of Development Studies, 54(10), 1792–1818. https://doi.org/10.1080/00220388.2017.1327663

Arnould, E. J., Plastina, A., & Ball, D. (2009). Does Fairtrade deliver on its core value proposition? Effects on income, educational attainment, and health in three countries. Journal of Public Policy & Marketing, 28(2), 186–201.

Blackman, A., & Naranjo, M. A. (2012). Does eco-certification have environmental benefits? Organic coffee in Costa Rica. Ecological Economics, 83, 58–66. https://doi.org/10.1016/j.ecolecon.2012.08.001

Blackman, A., & Rivera, J. (2011). Producer-level benefits of sustainability certification. Conservation Biology: The Journal of the Society for Conservation Biology, 25(6), 1176–1185. https://doi.org/10.1111/j.1523-1739.2011.01774.x

Bolwig, S., Gibbon, P., & Jones, S. (2009). The Economics of Smallholder Organic Contract Farming in Tropical Africa. World Development, 37(6), 1094–1104. https://doi.org/10.1016/j.worlddev.2008.09.012

Butsic, V., Lewis, D. J., Radeloff, V. C., Baumann, M., & Kuemmerle, T. (2017). Quasi-experimental methods enable stronger inferences from observational data in ecology. Basic and Applied Ecology, 19, 1–10. https://doi.org/10.1016/j.baae.2017.01.005

Breslow, S. J., Allen, M., Holstein, D., Sojka, B., Barnea, R., Basurto, X., Carothers, C., Charnley, S., Coulthard, S., Dolšak, N., Donatuto, J., García-Quijano, C., Hicks, C. C., Levine, A., Mascia, M. B., Norman, K., Poe, M., Satterfield, T., St. Martin, K., & Levin, P. S. (2017). Evaluating indicators of human well-being for ecosystem-based management. Ecosystem Health and Sustainability, 3(12), 1–18. https://doi.org/10.1080/20964129.2017.1411767

Breslow, S. J., Sojka, B., Barnea, R., Basurto, X., Carothers, C., Charnley, S., Coulthard, S., Dolšak, N., Donatuto, J., García-Quijano, C., Hicks, C. C., Levine, A., Mascia, M. B., Norman, K., Poe, M., Satterfield, T., Martin, K. St., & Levin, P. S. (2016). Conceptualizing and operationalizing human wellbeing for ecosystem assessment and management. Environmental Science & Policy, 66, 250–259. https://doi.org/10.1016/j.envsci.2016.06.023

Carlson, K. M., Heilmayr, R., Gibbs, H. K., Noojipady, P., Burns, D. N., Morton, D. C., Walker, N. F., Paoli, G. D., & Kremen, C. (2018). Effect of oil palm sustainability certification on deforestation and fire in Indonesia. Proceedings of the National Academy of Sciences, 115(1), 121–126. https://doi.org/10.1073/pnas.1704728114

Cattau, M. E., Marlier, M. E., & DeFries, R. (2016). Effectiveness of Roundtable on Sustainable Palm Oil (RSPO) for reducing fires on oil palm concessions in Indonesia from 2012 to 2015. Environmental Research Letters, 11(10), 105007. https://doi.org/10.1088/1748-9326/11/10/105007

Caudill, S. A., & Rice, R. A. (2016). Do Bird Friendly® Coffee Criteria Benefit Mammals? Assessment of Mammal Diversity in Chiapas, Mexico. PLOS ONE, 11(11), e0165662. https://doi.org/10.1371/journal.pone.0165662

Chiputwa, B., & Qaim, M. (2016). Sustainability Standards, Gender, and Nutrition among Smallholder Farmers in Uganda. The Journal of Development Studies, 52(9), 1241–1257. https://doi.org/10.1080/00220388.2016.1156090

Chiputwa, B., Spielman, D. J., & Qaim, M. (2015). Food Standards, Certification, and Poverty among Coffee Farmers in Uganda. World Development, 66, 400–412. https://doi.org/10.1016/j.worlddev.2014.09.006

Colen, L., Maertens, M., & Swinnen, J. (2012). Private Standards, Trade and Poverty: GlobalGAP and Horticultural Employment in Senegal. The World Economy, 35(8), 1073–1088. https://doi.org/10.1111/j.1467-9701.2012.01463.x

COSA: Committee on Sustainability Assessment. Indicator Library: Master List. (2020). https://thecosa.org/master-list/

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. https://doi.org/10.1016/j.agee.2013.05.003

DeFries, R. S., Fanzo, J., Mondal, P., Remans, R., & Wood, S. A. (2017). Is voluntary certification of tropical agricultural commodities achieving sustainability goals for small-scale producers? A review of the evidence. Environmental Research Letters, 12(3), 033001. https://doi.org/10.1088/1748-9326/aa625e

Dietz, T., Grabs, J., & Chong, A. E. (2018). Mainstreamed voluntary sustainability standards and their effectiveness: Evidence from the Honduran coffee sector. Regulation & Governance. https://doi.org/10.1111/rego.12239

Dietz, T., Chong, A. E., Grabs, J., & Kilian, B. (2020). How Effective is Multiple Certification in Improving the Economic Conditions of Smallholder Farmers? Evidence from an Impact Evaluation in Colombia's Coffee Belt. The Journal of Development Studies, 56(6), 1141–1160. https://doi.org/10.1080/00220388.2019.1632433

Doanh, N. K., Thuong, N. T. T., & Heo, Y. (2018). Impact of Conversion to Organic Tea Cultivation on Household Income in the Mountainous Areas of Northern Vietnam. Sustainability, 10(12), 4475. https://doi.org/10.3390/su10124475

Donovan, J., & Blare, T. (2018). Evaluation of UTZ certification focused on coffee businesses in Guatemala, Honduras and Nicaragua. Nairobi: World Agroforestry Centre.

Elbers, W., van Rijsbergen, B., Bagamba, F., & Hoebink, P. (2014). Chapter 2 The impact of Utz certification on smallholder farmers in Uganda. In Coffee certification in East Africa: impact on farms, families and cooperatives (Vol. 1–0, pp. 53–82). Wageningen Academic Publishers. https://doi.org/10.3920/978-90-8686-805-6_2

Ferraro, P. J. (2009). Counterfactual thinking and impact evaluation in environmental policy. New Directions for Evaluation, 2009(122), 75–84. https://doi.org/10.1002/ev.297

Filho, M. A. R., Andow, D. A., Carneiro, R. G., Lorena, D. R., Sujii, E. R., & Alves, R. T. (2019). Economic and productivity incentives to produce organically in Brazil: Evidence from strawberry production in the Federal District. Renewable Agriculture and Food Systems, 34(2), 155–168. https://doi.org/10.1017/S1742170517000412

Ferraro, P. J., & Hanauer, M. M. (2014). Advances in Measuring the Environmental and Social Impacts of Environmental Programs. Annual Review of Environment and Resources, 39(1), 495–517. https://doi.org/10.1146/annurev-environ-101813-013230

Foley, J. A., Ramankutty, N., Brauman, K. A., Cassidy, E. S., Gerber, J. S., Johnston, M., Mueller, N. D., O'Connell, C., Ray, D. K., West, P. C., Balzer, C., Bennett, E. M., Carpenter, S. R., Hill, J., Monfreda, C., Polasky, S., Rockström, J., Sheehan, J., Siebert, S., … Zaks, D. P. M. (2011). Solutions for a cultivated planet. Nature, 478(7369), 337–342. https://doi.org/10.1038/nature10452

Fort, R., & Ruben, R. (2009). The impact of Fairtrade on banana producers in northern Peru. In 2009 Conference, August 16-22, 2009, Beijing, China (No. 50964; 2009 Conference, August 16-22, 2009, Beijing, China). International Association of Agricultural Economists. https://ideas.repec.org/p/ags/iaae09/50964.html

Froehlich, A. G., Melo, A. S. S. A., & Sampaio, B. (2018). Comparing the Profitability of Organic and Conventional Production in Family Farming: Empirical Evidence From Brazil. Ecological Economics, 150(C), 307–314.

Garibaldi, L. A., Gemmill-Herren, B., D'Annolfo, R., Graeub, B. E., Cunningham, S. A., & Breeze, T. D. (2017). Farming Approaches for Greater Biodiversity, Livelihoods, and Food Security. Trends in Ecology & Evolution, 32(1), 68–80. https://doi.org/10.1016/j.tree.2016.10.001

Haggar, J., Soto, G., Casanoves, F., & Virginio, E. de M. (2017). Environmental-economic benefits and trade-offs on sustainably certified coffee farms. Ecological Indicators, 79, 330–337. https://doi.org/10.1016/j.ecolind.2017.04.023

Ibanez, M., & Blackman, A. (2016). Is Eco-Certification a Win–Win for Developing Country Agriculture? Organic Coffee Certification in Colombia. World Development, 82, 14–27. https://doi.org/10.1016/j.worlddev.2016.01.004

IISD: International Institute for Sustainable Development. (2015). Voluntary Sustainability Standards. IISD: International Institute for Sustainable Development. https://www.iisd.org/topic/standards

Ingram, V., Van Rijn, F., Waarts, Y., & Gilhuis, H. (2018). The Impacts of Cocoa Sustainability Initiatives in West Africa. Sustainability, 10(11), 4249. https://doi.org/10.3390/su10114249

ISEAL Alliance. Multiple Certification in Coffee & Cocoa. (2018). London: UK. https://www.isealalliance.org/sites/default/files/resource/2019-02/DIPI_Multiple_Certification_Analysis_Final_Oct%202018.pdf

ISEAL Alliance. Introduction to Sustainability Benchmarking. (2019). https://www.isealalliance.org/sites/default/files/resource/2019-05/Introduction_to_sustainability_benchmarking-ISEAL-2018.pdf

ITC (International Trade Centre). Standards Map. (2015). http://standardsmap.org/identify2.aspx

Jena, P. R., Chichaibelu, B. B., Stellmacher, T., & Grote, U. (2012). The impact of coffee certification on small-scale producers' livelihoods: a case study from the Jimma Zone, Ethiopia. Agricultural Economics, 43(4), 429–440. https://doi.org/10.1111/j.1574-0862.2012.00594.x

Jena, P., Stellmacher, T., & Grote, U. (2015). Can coffee certification schemes increase incomes of smallholder farmers? Evidence from Jinotega, Nicaragua. Environment, Development and Sustainability, 19. https://doi.org/10.1007/s10668-015-9732-0

Jena, P., & Grote, U. (2017). Fairtrade Certification and Livelihood Impacts on Small-scale Coffee Producers in a Tribal Community of India. Applied Economic Perspectives and Policy, 39(1), 87–110. https://doi.org/10.1093/aepp/ppw006

Jha, S., Bacon, C. M., Philpott, S. M., Ernesto Méndez, V., Läderach, P., & Rice, R. A. (2014). Shade Coffee: Update on a Disappearing Refuge for Biodiversity. BioScience, 64(5), 416–428. https://doi.org/10.1093/biosci/biu038

Jones, S., & Gibbon, P. (2011). Developing Agricultural Markets in Sub-Saharan Africa: Organic Cocoa in Rural Uganda. The Journal of Development Studies, 47(10), 1595–1618. https://doi.org/10.1080/00220388.2011.579107

Karki, S. K., Jena, P. R., & Grote, U. (2016). Fairtrade Certification and Livelihoods: A Panel Data Analysis of Coffee-growing Households in India. https://doi.org/10.15488/1712

Kleemann, L., & Abdulai, A. (2013). Organic certification, agro-ecological practices and return on investment: Evidence from pineapple producers in Ghana. Ecological Economics, 93, 330–341. https://doi.org/10.1016/j.ecolecon.2013.06.017

Komives, K., Arton, A., Baker, E., Kennedy, E., Longo, C., Newsom, D., Pfaff, A., & Romero, C. (2018). Conservation impacts of voluntary sustainability standards: How has our understanding changed since the 2012 publication of "Toward sustainability: The roles and limitations of certification"? Washington DC: Meridian Institute. Available at merid.org/content/projects/supply_chain_sustainability_research_fund.

Last, L., Arndorfer, M., Balázs, K., Dennis, P., Dyman, T., Fjellstad, W., Friedel, J. K., Herzog, F., Jeanneret, P., Lüscher, G., Moreno, G., Kwikiriza, N., Gomiero, T., Paoletti, M. G., Pointereau, P., Sarthou, J.-P., Stoyanova, S., Wolfrum, S., & Kölliker, R. (2014). Indicators for the on-farm assessment of crop cultivar and livestock breed diversity: a survey-based participatory approach. Biodiversity and Conservation, 23(12), 3051–3071. https://doi.org/10.1007/s10531-014-0763-x

Latruffe, L., Diazabakana, A., Bockstaller, C., Desjeux, Y., Finn, J., Kelly, E., Ryan, M., & Uthes, S. (2016). Measurement of sustainability in agriculture: a review of indicators. Studies in Agricultural Economics, 118(3), 123–130. https://doi.org/10.7896/j.1624

Margoluis, R., Stem, C., Salafsky, N., & Brown, M. (2009). Design alternatives for evaluating the impact of conservation projects. New Directions for Evaluation, 2009(122), 85–96. https://doi.org/10.1002/ev.298

Mascia, M. B., Pailler, S., Thieme, M. L., Rowe, A., Bottrill, M. C., Danielsen, F., Geldmann, J., Naidoo, R., Pullin, A. S., & Burgess, N. D. (2014). Commonalities and complementarities among approaches to conservation monitoring and evaluation. Biological Conservation, 169, 258–267. https://doi.org/10.1016/j.biocon.2013.11.017

Meemken, E.-M., & Qaim, M. (2018). Can private food standards promote gender equality in the small farm sector? Journal of Rural Studies, 58, 39–51. https://doi.org/10.1016/j.jrurstud.2017.12.030

Meemken, E.-M., Spielman, D. J., & Qaim, M. (2017). Trading off nutrition and education? A panel data analysis of the dissimilar welfare effects of Organic and Fairtrade standards. Food Policy, 71, 74–85. https://doi.org/10.1016/j.foodpol.2017.07.010

Meemken, Eva-Marie. 2020. "Do Smallholder Farmers Benefit from Sustainability Standards? A Systematic Review and Meta-Analysis." Global Food Security 26 (September): 100373. https://doi.org/10.1016/j.gfs.2020.100373.

Milder, J. C., Arbuthnot, M., Blackman, A., Brooks, S. E., Giovannucci, D., Gross, L., Kennedy, E. T., Komives, K., Lambin, E. F., Lee, A., Meyer, D., Newton, P., Phalan, B., Schroth, G., Semroc, B., Rikxoort, H. V., & Zrust, M. (2015). An agenda for assessing and improving conservation impacts of sustainability standards in tropical agriculture. Conservation Biology, 29(2), 309–320. https://doi.org/10.1111/cobi.12411

Minten, B., Dereje, M., Engida, E., & Tamru, S. (2018). Tracking the Quality Premium of Certified Coffee: Evidence from Ethiopia. World Development, 101(C), 119–132.

Mitiku, F., De Mey, Y., Nyssen, J., & Maertens, M. (2017). Do Private Sustainability Standards Contribute to Income Growth and Poverty Alleviation? A Comparison of Different Coffee Certification Schemes in Ethiopia. Sustainability, 9(2), 246. https://doi.org/10.3390/su9020246

Mitiku, F., Nyssen, J., & Maertens, M. (2018). Certification of Semi-forest Coffee as a Land-sharing Strategy in Ethiopia. Ecological Economics, 145, 194–204. https://doi.org/10.1016/j.ecolecon.2017.09.008

Moher D, Liberati A, Tetzlaff J, Altman DG, The PRISMA Group (2009). *P*referred *R*eporting *I*tems for *S*ystematic Reviews and *M*eta-*A*nalyses: The PRISMA Statement. PLoS Med 6(7): e1000097. doi:10.1371/journal.pmed1000097

Morgans, C. L., Meijaard, E., Santika, T., Law, E., Budiharta, S., Ancrenaz, M., & Wilson, K. A. (2018). Evaluating the effectiveness of palm oil certification in delivering multiple sustainability objectives. Environmental Research Letters, 13(6), 064032. https://doi.org/10.1088/1748-9326/aac6f4

Naidoo, R., Gerkey, D., Hole, D., Pfaff, A., Ellis, A. M., Golden, C. D., Herrera, D., Johnson, K., Mulligan, M., Ricketts, T. H., & Fisher, B. (2019). Evaluating the impacts of protected areas on human well-being across the developing world. Science Advances, 5(4), eaav3006. https://doi.org/10.1126/sciadv.aav3006

O'Rourke, D. (2014). The science of sustainable supply chains. Science, 344(6188), 1124–1127. https://doi.org/10.1126/science.1248526

Oya, C., Schaefer, F., & Skalidou, D. (2018). The effectiveness of agricultural certification in developing countries: A systematic review. World Development, 112, 282–312. https://doi.org/10.1016/j.worlddev.2018.08.001

Petrokofsky, G., & Jennings, S. (2018). The effectiveness of standards in driving adoption of sustainability practices: A State of Knowledge Review. 3Keel and ISEAL Alliance.

Potts, J., Lynch, M., Wilkings, A., Huppe, G., Cunningham, M., & Voora, V. (2014). The state of sustainability initiatives review 2014: Sustainability and transparency. International Institute for Sustainable Development and the International Institute for Environment and Development.

Qiao, Y., Martin, F., Cook, S., He, X., Halberg, N., Scott, S., & Pan, X. (2018). Certified Organic Agriculture as an Alternative Livelihood Strategy for Small-scale Farmers in China: A Case Study in Wanzai County, Jiangxi Province. Ecological Economics, 145(C), 301–307.

Rainforest Alliance. (2017). Rainforest Alliance - Sustainable Agriculture Standard Version 1.2. Red de Agricultural Sostenible, A.C.(Sustainable Agriculture Network). https://www.rainforest-alliance.org/business/wp-content/uploads/2017/11/03_rainforest-alliance-sustainable-agriculture-standard_en.pdf

Rasmussen, L. V., Coolsaet, B., Martin, A., Mertz, O., Pascual, U., Corbera, E., Dawson, N., Fisher, J. A., Franks, P., & Ryan, C. M. (2018). Social-ecological outcomes of agricultural intensification. Nature Sustainability, 1(6), 275–282. https://doi.org/10.1038/s41893-018-0070-8

R Core Team (2013). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-project.org/.

Roundtable on Sustainable Biomaterials. 2018. "Understanding Our Footprint: Outcomes of RSB's Monitoring & Evaluation System in 2018." Roundtable on Sustainable Biomaterials. https://rsb.org/wp-content/uploads/2019/06/RSB-Monitoring-Evaluation-System-Report-2018-Draft-2.pdf.

Ruben, R., Fort, R., & Zúñiga-Arias, G. (2009). Measuring the impact of Fairtrade on development. Development in Practice, 19(6), 777–788. https://doi.org/10.1080/09614520903027049

Ruben, R., & Fort, R. (2012). The Impact of Fairtrade Certification for Coffee Farmers in Peru. World Development, 40(3), 570–582.

Ruben, R., & Zuniga, G. (2011). How standards compete: comparative impact of coffee certification schemes in Northern Nicaragua. Supply Chain Management: An International Journal, 16(2), 98–109. https://doi.org/10.1108/13598541111115356

Rueda, X., & Lambin, E. (2013). Responding to Globalization: Impacts of Certification on Colombian Small-Scale Coffee Growers. Ecology and Society, 18(3). https://doi.org/10.5751/ES-05595-180321

Rueda, X., Thomas, N. E., & Lambin, E. F. (2015). Eco-certification and coffee cultivation enhance tree cover and forest connectivity in the Colombian coffee landscapes. Regional Environmental Change, 15(1), 25–33. https://doi.org/10.1007/s10113-014-0607-y

SAI Platform. (2014). Sustainability Performance Assessment Version 2.0. https://saiplatform.org/uploads/SPA%20Guidelines%202%200.pdf

Santika, Truly, Kerrie A. Wilson, Elizabeth A. Law, Freya A. V. St John, Kimberly M. Carlson, Holly Gibbs, Courtney L. Morgans, Marc Ancrenaz, Erik Meijaard, and Matthew J. Struebig. 2020. "Impact of Palm Oil Sustainability Certification on Village Well-Being and Poverty in Indonesia." *Nature Sustainability*, November, 1–11. https://doi.org/10.1038/s41893-020-00630-1.

Schoonhoven-Speijer, M., & Ruben, R. (2014). Chapter 5 Maintaining sustainable livelihoods: effects of Utz certification on market access, risk reduction and livelihood strategies of Kenyan coffee farmers. In Coffee certification in East Africa: impact on farms, families and cooperatives (Vol. 1–0, pp. 149–174). Wageningen Academic Publishers. https://doi.org/10.3920/978-90-8686-805-6_5

Sellare, Jorge, Eva-Marie Meemken, and Matin Qaim. 2020. "Fairtrade, Agrochemical Input Use, and Effects on Human Health and the Environment." *Ecological Economics* 176 (October): 106718. https://doi.org/10.1016/j.ecolecon.2020.106718.

Siddaway, A. P., Wood, A. M., & Hedges, L. V. (2019). How to Do a Systematic Review: A Best Practice Guide for Conducting and Reporting Narrative Reviews, Meta-Analyses, and Meta-Syntheses. Annual Review of Psychology, 70(1), 747–770. https://doi.org/10.1146/annurev-psych-010418-102803

Smith, W. K., Nelson, E., Johnson, J. A., Polasky, S., Milder, J. C., Gerber, J. S., West, P. C., Siebert, S., Brauman, K. A., Carlson, K. M., Arbuthnot, M., Rozza, J. P., & Pennington, D. N. (2019). Voluntary sustainability standards could significantly reduce detrimental impacts of global agriculture. Proceedings of the National Academy of Sciences, 116(6), 2130–2137. https://doi.org/10.1073/pnas.1707812116

Takahashi, R., & Todo, Y. (2013). The impact of a shade coffee certification program on forest conservation: a case study from a wild coffee forest in Ethiopia. Journal of Environmental Management, 130, 48–54. https://doi.org/10.1016/j.jenvman.2013.08.025

Takahashi, R., & Todo, Y. (2014). The impact of a shade coffee certification program on forest conservation using remote sensing and household data. Environmental Impact Assessment Review, 44, 76–81. https://doi.org/10.1016/j.eiar.2013.10.002

Takahashi, R., & Todo, Y. (2017). Coffee Certification and Forest Quality: Evidence from a Wild Coffee Forest in Ethiopia. World Development, 92, 158–166. https://doi.org/10.1016/j.worlddev.2016.12.001
Targetti, S., Herzog, F., Geijzendorffer, I. R., Wolfrum, S., Arndorfer, M., Balàzs, K., Choisis, J. P., Dennis, P., Eiter, S., Fjellstad, W., Friedel, J. K., Jeanneret, P., Jongman, R. H. G., Kainz, M., Luescher, G., Moreno, G., Zanetti, T., Sarthou, J. P., Stoyanova, S., … Viaggi, D. (2014). Estimating the cost of different strategies for measuring farmland biodiversity: Evidence from a Europe-wide field evaluation. Ecological Indicators, 45, 434–443. https://doi.org/10.1016/j.ecolind.2014.04.050
Tayleur, C., Balmford, A., Buchanan, G. M., Butchart, S. H. M., Ducharme, H., Green, R. E., Milder, J. C., Sanderson, F. J., Thomas, D. H. L., Vickery, J., & Phalan, B. (2017a). Global Coverage of Agricultural Sustainability Standards, and Their Role in Conserving Biodiversity: Certification standards and biodiversity. Conservation Letters, 10(5), 610–618. https://doi.org/10.1111/conl.12314. License: https://creativecommons.org/licenses/by/4.0/
Tayleur, Catherine, Juliet Vickery, Stuart Butchart, Christine Corlet Walker, Graeme Buchanan, Fiona Sanderson, Jeffrey Milder, et al. 2017b. "GIS Data for: Where Are Commodity Crops Certified, and What Does It Mean for Conservation and Poverty Alleviation?" 2 (October). https://doi.org/10.17632/mpdf6ytswm.2. License: https://creativecommons.org/licenses/by-nc/3.0/
Tayleur, C., Balmford, A., Buchanan, G. M., Butchart, S. H. M., Corlet Walker, C., Ducharme, H., Green, R. E., Milder, J. C., Sanderson, F. J., Thomas, D. H. L., Tracewski, L., Vickery, J., & Phalan, B. (2018). Where are commodity crops certified, and what does it mean for conservation and poverty alleviation? Biological Conservation, 217, 36–46. https://doi.org/10.1016/j.biocon.2017.09.024
The World Bank. World Bank Country and Lending Groups. (2020). https://datahelpdesk.worldbank.org/knowledgebase/articles/906519
Tran, D., & Goto, D. (2019). Impacts of sustainability certification on farm income: Evidence from small-scale specialty green tea farmers in Vietnam. Food Policy, 83, 70–82. https://doi.org/10.1016/j.foodpol.2018.11.006
Tscharntke, T., Milder, J. C., Schroth, G., Clough, Y., DeClerck, F., Waldron, A., Rice, R., & Ghazoul, J. (2015). Conserving Biodiversity Through Certification of Tropical Agroforestry Crops at Local and Landscape Scales. Conservation Letters, 8(1), 14–23. https://doi.org/10.1111/conl.12110
Vanderhaegen, K., Akoyi, K. T., Dekoninck, W., Jocqué, R., Muys, B., Verbist, B., & Maertens, M. (2018). Do private coffee standards 'walk the talk' in improving socio-economic and environmental sustainability? Global Environmental Change, 51, 1–9. https://doi.org/10.1016/j.gloenvcha.2018.04.014
van Rijsbergen, B., Elbers, W., Ruben, R., & Njuguna, S. N. (2016). The Ambivalent Impact of Coffee Certification on Farmers' Welfare: A Matched Panel Approach for Cooperatives in Central Kenya. World Development, 77, 277–292. https://doi.org/10.1016/j.worlddev.2015.08.021
Vitousek, P. M., Naylor, R., Crews, T., David, M. B., Drinkwater, L. E., Holland, E., Johnes, P. J., Katzenberger, J., Martinelli, L. A., Matson, P. A., Nziguheba, G., Ojima, D., Palm, C. A., Robertson, G. P., Sanchez, P. A., Townsend, A. R., & Zhang, F. S. (2009). Nutrient

Imbalances in Agricultural Development. Science, 324(5934), 1519–1520. https://doi.org/10.1126/science.1170261
Weber, J. G. (2011). How much more do growers receive for Fairtrade-organic coffee? Food Policy, 36(5), 678–685. https://doi.org/10.1016/j.foodpol.2011.05.007
Willer, H., Sampson, G., Voora, V., Dang, D., & Lernoud, J. (2019). State of Sustainable Markets 2019: Statistics and Emerging Trends. International Trade Center. https://www.deslibris.ca/ID/10102592
Zulfiqar, F., & Thapa, G. B. (2016). Is 'Better cotton' better than conventional cotton in terms of input use efficiency and financial performance? Land Use Policy, 52, 136–143. https://doi.org/10.1016/j.landusepol.2015.12.013

Chapter 3: The poverty and forest cover impacts of sustainable development projects in Lao PDR from 2005 to 2015[18]

3.1 Introduction

Environmental problems are, on the whole, rapidly worsening, with serious implications for people, wildlife, and ecosystems (Steffen et al. 2015). Considering such challenges, it is increasingly important to understand the effectiveness of sustainability interventions. Despite this, quasi-experimental approaches to infer causal effects are historically under-utilized in the conservation realm (Fisher et al. 2014, Mascia et al. 2014). This topic is receiving increasing attention vis-à-vis land-based climate change mitigation and deforestation reduction targets (S. Roe et al. 2021, Griscom et al. 2017, Caplow et al. 2011, Desbureaux 2021). In this study, we build on recent advancements in causal inference and environmental programs to assess the impact of voluntary sustainability programs in Laos (here referred to as 'sustainable development projects' or SDPs). Our main research question is: *To what extent did voluntary SDPs in Laos affect poverty and forest cover between 2005 and 2015?* We evaluate SDPs implemented between 2007 and 2013 through an inverse probability-of-treatment weighted regression (IPWR).

Laos presents an important case study to evaluate SDP impact. Laos has one of the highest proportional national forest covers globally and set forth the National Forest Strategy to increase forest cover to 70% by 2020 (FAO 2016, Smith et al. 2021). Despite this, the country has faced increasing deforestation pressure since the 1980s (Phompila et al. 2017). Deforestation

[18] This manuscript will be submitted to *Proceedings of the National Academy of Sciences* in May 2022. Co-authors on this chapter include Dr. Peter M. Steiner, Jeff Sauer, and Dr. Peter Potapov.

drivers include logging, mining, hydropower development, the growth of large-scale plantations, and shifting cultivation (World Bank Group and Climate Investment Funds 2019, Zeng et al. 2018, Boutthavong, Hyakumura, and Ehara 2017). Furthermore, although Laos achieved significant poverty reduction over the last few decades following economic reforms, the poverty rate remained 18% in 2019 (Lao Statistics Bureau and World Bank 2020, Bader et al. 2017). In addition to environmental consequences, deforestation also negatively impacts local communities who depend on forests for livelihoods and non-timber forest product (NTFP) access (Rigg 2006). Together, these challenges have prompted a variety of actions, including government regulation (e.g. logging bans and large-scale plantation moratoria) and voluntary programs via NGOs and international agencies.

This study examines a set of voluntary SDPs which aimed to reduce poverty, maintain forest cover, and/or protect wildlife habitat. These projects are representative of SDPs deployed in the Greater Mekong region more generally and focused on food security and livelihoods diversification; sustainable forest management; and sustainable resource management (see Appendix 2, Table A2.5). Food security and livelihoods projects involved irrigation, livestock, nutrition, and livelihoods development. Forest management SDPs incorporated both productive forestry (e.g., smallholder bamboo plantations) and forest protection, both within and outside of formal protected areas or national forests. Resource management projects focused primarily on habitat conservation.

Previous research underscores the context-dependence and effect heterogeneity of SDPs - which are influenced by the surrounding socioeconomic, market, and policy environment as well as the project's design (Miteva, Pattanayak, and Ferraro 2012). However, the literature suggests several trends in their impacts. For example, irrigation improvement projects have resulted in positive impacts both for poverty (Hussain and Hanjra 2004) and deforestation (Baehr, BenYishay, and Parks 2021, Shively and Pagiola 2004). Small-scale livestock diversification has also been found to result in poverty alleviation, with potential environmental trade-offs (Udo et al. 2011, Hao et al. 2015). Plantation forestry studies have shown mixed outcomes both for poverty and forests, and studies suggest that the type of forestry arrangement (e.g. community forestry, concession, smallholder production) significantly impacts success of the intervention (Phimmavong et al. 2019, Pirard, Dal Secco, and Warman 2016, Malkamäki et al. 2018, Andersson et al. 2016, van der Meer Simo 2020, Coleman et al. 2021). Challenges can include income foregone from resource extraction or agricultural livelihoods, or pre-existing land and income inequalities (van der Meer Simo 2020, Andersson et al. 2016). Previous analyses of protected areas have shown that habitat conservation can have positive effects on human well-being (Naidoo et al. 2019, Ferraro, Hanauer, and Sims 2011, Andam et al. 2010), as well as the potential for negative social impacts, including community displacement (Hajjar et al. 2021). However, a lack of more rigorous quasi-experimental studies still limits our ability to generalize conclusions regarding intervention effectiveness (Malkamäki et al. 2018, Caplow et al. 2011, McKinnon et al. 2016, Puri et al. 2017).

We hypothesize that SDPs focused on food security and livelihood diversification will have poverty-reducing effects, based on previous empirical research and SDP project evaluations during the study period (Newby et al. 2012, Hussain and Hanjra 2004, Udo et al. 2011, World Bank Group and Climate Investment Funds 2019, Coudray 2014, Asian Development Bank 2011). We also hypothesize that SDPs will have positive effects on forest cover maintenance, particularly sustainable forest and resource management projects (Honey-Rosés, Baylis, and Ramírez 2011). By evaluating poverty and forest cover simultaneously, we also assess potential co-benefits of SDPs for both outcomes. In Laos, population pressure and reduced access to land have caused shorter follow periods and increased land conversion from shifting cultivation; many communities practicing shifting cultivation continue to face high poverty rates and are dependent on the practice for food supply (Seidenberg, Mertz, and Kias 2003, World Bank Group and Climate Investment Funds 2019, Phompila et al. 2017, Rigg 2006). Therefore, alternative livelihood options for communities practicing shifting cultivation may bring dual benefits.

We use IPWR to evaluate the impact of SDPs implemented in 1,452 SDP villages, compared to 6,615 matched non-SDP villages. Our analysis combines inverse probability-of-treatment weights (IPTWs) from a propensity score (PS) model with an additional regression adjustment (Rosenbaum and Rubin 1983, Steiner and Cook 2013, Imbens and Wooldridge 2009). To account for potential confounding bias due to differential enrollment of villages in SDPs, we use IPTWs estimated from a set of 32 carefully chosen covariates in a weighted regression that in addition controls for the same covariate set. We evaluate the average treatment effect for the treated villages (ATT) for (1) forest cover (the village's proportional forested area in 2015) and (2) poverty outcomes (2015 % poverty headcount). Outcome variables are displayed in Figure 3.1.

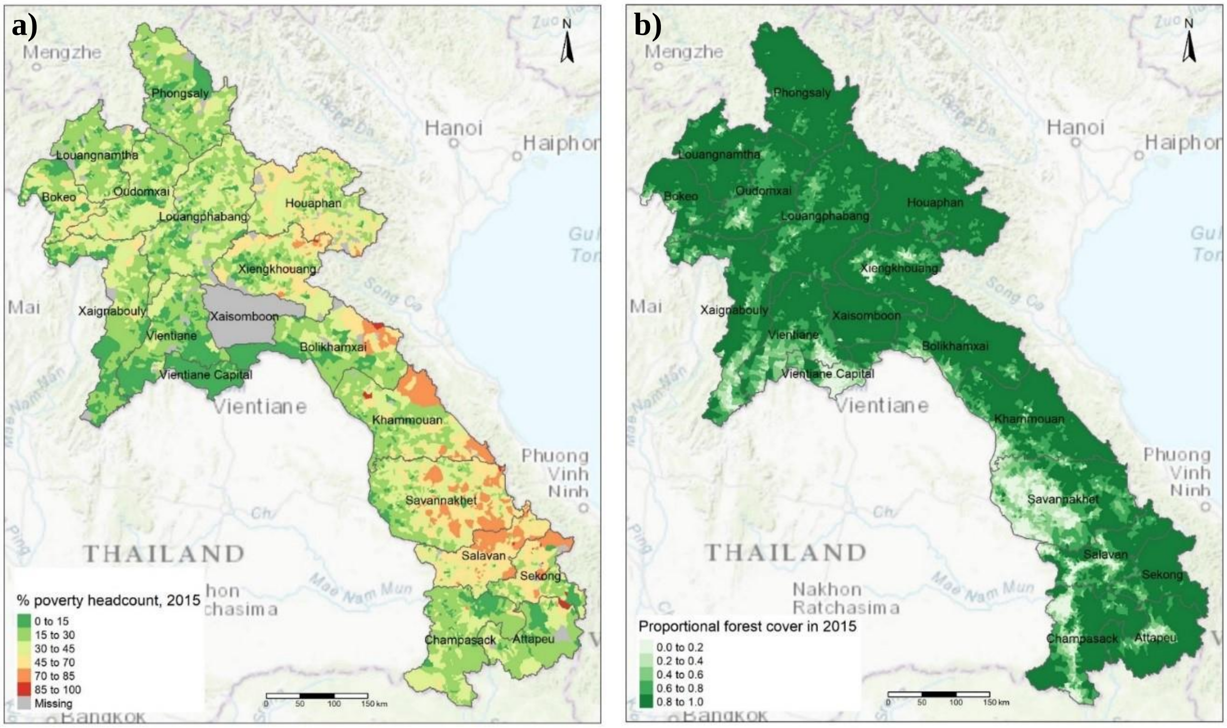


Figure 3.1: Map displaying (a) % poverty headcounts in Laos (2015) and (b) village-level proportional forest cover in 2015 for village land-use areas. Authors' elaboration based on (1) Esri Topographic map; (2) Laos village polygons (Epprecht et al. (2008), Messerli et al. (2008), and the 2015 Laos Population and Housing Census); (3) poverty headcounts from the 2015 Laos Population and Housing Census; and (4) forest cover data derived from the Landsat data archive, following Potapov et al. 2019 and Turubanova et al. 2018. Missing villages are those which could not be matched through the geospatial linking procedure described in Section A2.2.3., as well as missing data from Xaisomboon provinces.

The SDP and non-SDP villages differ due to baseline characteristics which simultaneously influenced their ability and willingness to engage with SDPs (e.g. socioeconomic status) and our sustainability outcomes of interest. Therefore, we rely on causal graphs, also known as directed acyclic graphs (DAGs, Pearl 2009), to determine a bias-removing adjustment set of covariates. DAGs have previously appeared in the environmental evaluation literature, although their use is not widespread (for examples see Ferraro and Hanauer 2014, Adams et al. 2015, Ferraro and Simorangkir 2020, Ferraro, Sanchirico, and Smith 2019). The DAGs, outlined in Appendix A2.1, represent subject matter theory regarding drivers of SDP interventions, poverty, and forest cover change. They allow us to assess whether the causal effect is identified given the observed set of covariates, and whether any remaining bias might seriously invalidate the study's conclusions (Elwert 2013, Rohrer 2018). Furthermore, due to variations in SDP design in Laos during the study period, we develop an SDP typology to differentiate among them and identify potential heterogeneities in sustainability impacts (see Section 3.4 and Appendix A2.3). This study makes three main contributions to the literature on impact evaluation of voluntary SDPs – first, it represents a novel use of DAGs to guide covariate selection and model design; second, it illustrates the value of impact evaluations in elucidating real-world successes and challenges of sustainability programs, including potential co-benefits for both poverty reduction and forest protection; and third, it provides lessons learned for sustainable development in the Greater Mekong, and more generally for 'nature-based interventions'.

3.2 Results

3.2.1 Sustainable development project villages

We leverage a database of SDPs in Laos from 2005-2015 provided by the Center for Development and Environment of the University of Bern (CDE) and linked to national village land-use area polygons (see Figure 3.2 and Appendix A2.2.3). SDP project start dates ranged from 2007-2013, apart from one project initiated in 2004. The majority (84%) of projects started between 2008-2010, with an average projection duration of 3.7 years. We assess sustainability impacts for four SDP groups – (1) all SDPs together, (2) food security and livelihoods projects, (3) sustainable forest management projects, and (4) sustainable resource management projects. Project types were delineated based on theories of change for affecting poverty and forest cover outcomes (Appendix A2.3). Food security and livelihoods projects encompassed irrigation improvements; livestock development (e.g. goats, pigs, chickens); rice and other agriculture;

nutrition; poverty alleviation; domesticated NTFP production; and agro-enterprise development. Forest management projects included protected area management; general forest conservation and avoidance of deforestation/degradation; and productive forestry (bamboo, teak). Resource management projects incorporated ecosystem and habitat conservation (wetlands, habitat for tigers, bears, deer, and crocodiles), and projects focused on multifunctional land-use and agro-biodiversity. SDPs were funded by more than 30 different donors, including the European Commission and FAO. Notably, long-term trends in mean forest cover among the SDP villages and control group exhibit some differences (Figure 3.3). All exhibit a downward trend in mean forest cover, although forest management SDP villages see the largest dip in 2015.

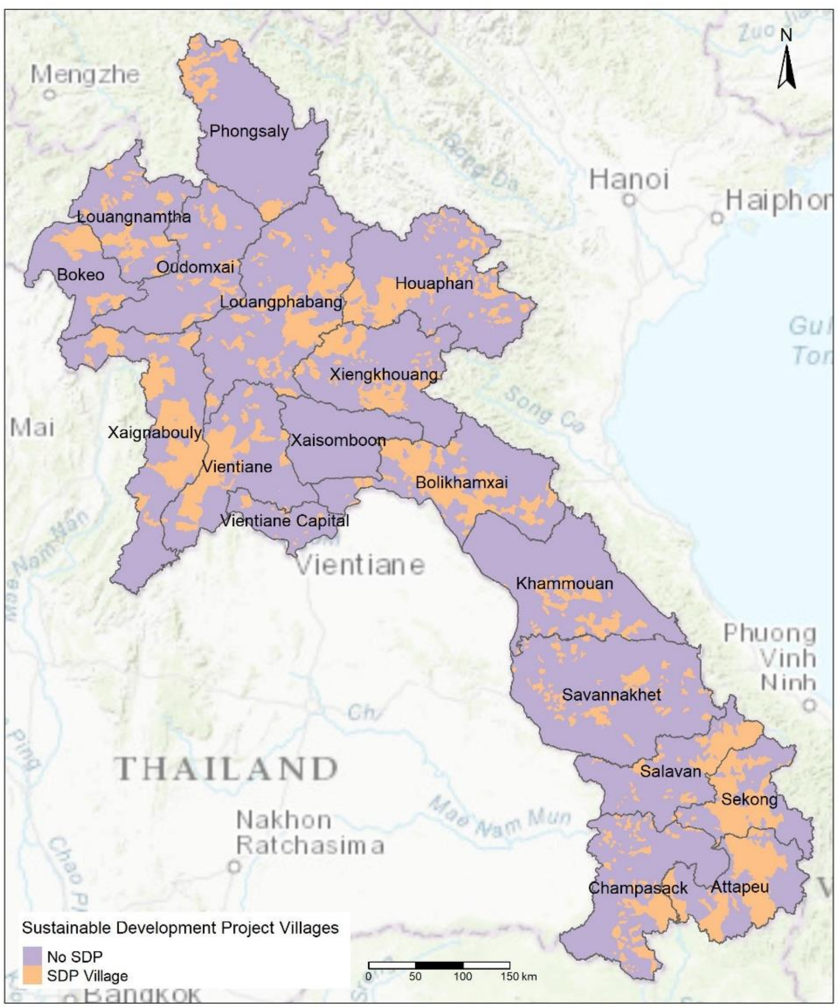


Figure 3.2: SDPs in Laos between 2005-2015 included in this study. SDPs were fairly distributed throughout the country but were most common in Louangphabang, Houaphan, and Xiengkhouang. Source: Authors' elaboration based on (1) Esri Topographic map; (2) Laos village polygons (Epprecht et al. (2008), Messerli et al. (2008), and the 2015 Laos Population and Housing Census); and (3) SDP database provided by CDE.

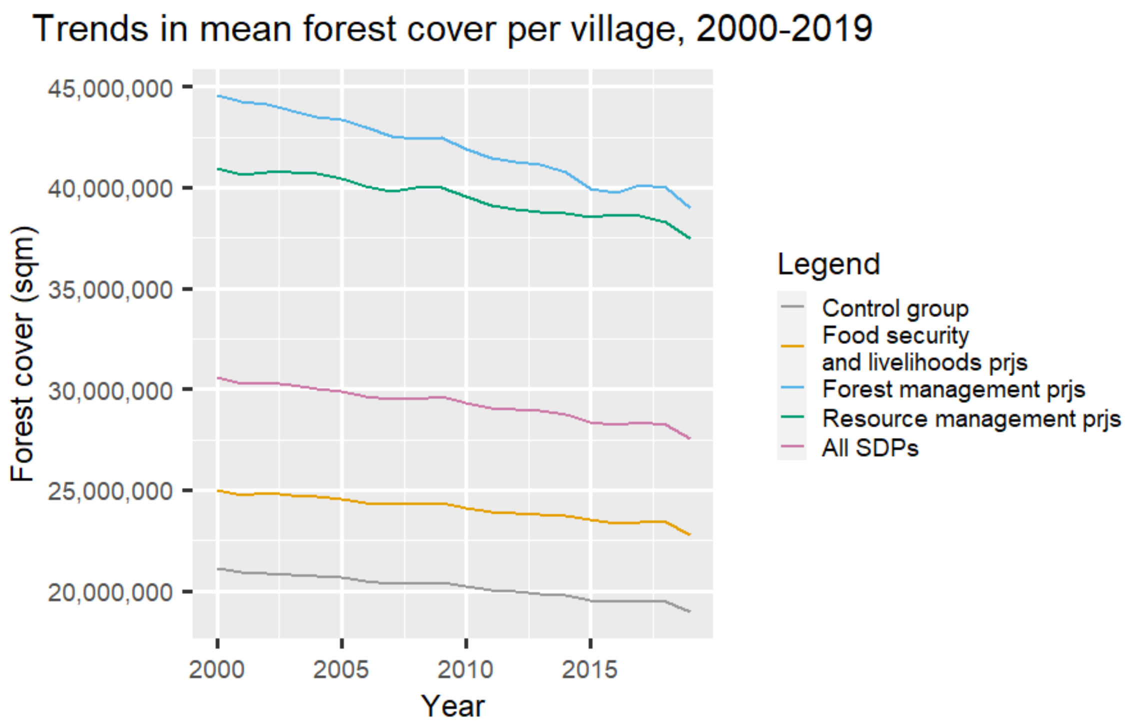


Figure 3.3: Annual mean forest cover by SDP project type.

3.2.2 SDP selection bias and inverse probability-of-treatment weights

SDP and non-SDP villages exhibited several important differences in baseline characteristics. For example, SDP villages tended to have higher baseline forest cover and elevation, a lower proportion of the population from the Lao majority ethno-linguistic group, and higher baseline poverty as compared to non-SDP villages (see Table 3.1). SDP villages also tended to be closer to protected areas and have a higher proportion of households with farmland. This comports with our DAGs (Appendix A2.1). We assessed the quality of the IPTWs using Love plots and kernel density plots illustrating pre-and post-matching statistics. The IPTWs balance the covariate distributions across all modeled SDP types (see Figure 3.4, and Appendix A2.5.2).

Table 3.1: Illustrative baseline (2005) characteristics for SDPs and non-SDPs (mean).

	All SDP villages (n=1,452)	Food security and livelihoods SDPs (n= 829)	Sustainable forest management SDPs (n= 166)	Sustainable resource management (n= 375)	Non-SDP villages (n=6,615)
Poverty (% headcount)	44.6%	44.5%	45.3%	46.5%	39.5%
Proportional forest cover	.75	.71	.82	.78	.65
Elevation (m)	627	647	520	668	538
Population Lao ethno-linguistic group (%)	21.5%	24.2%	15.8%	21.3%	32.5%

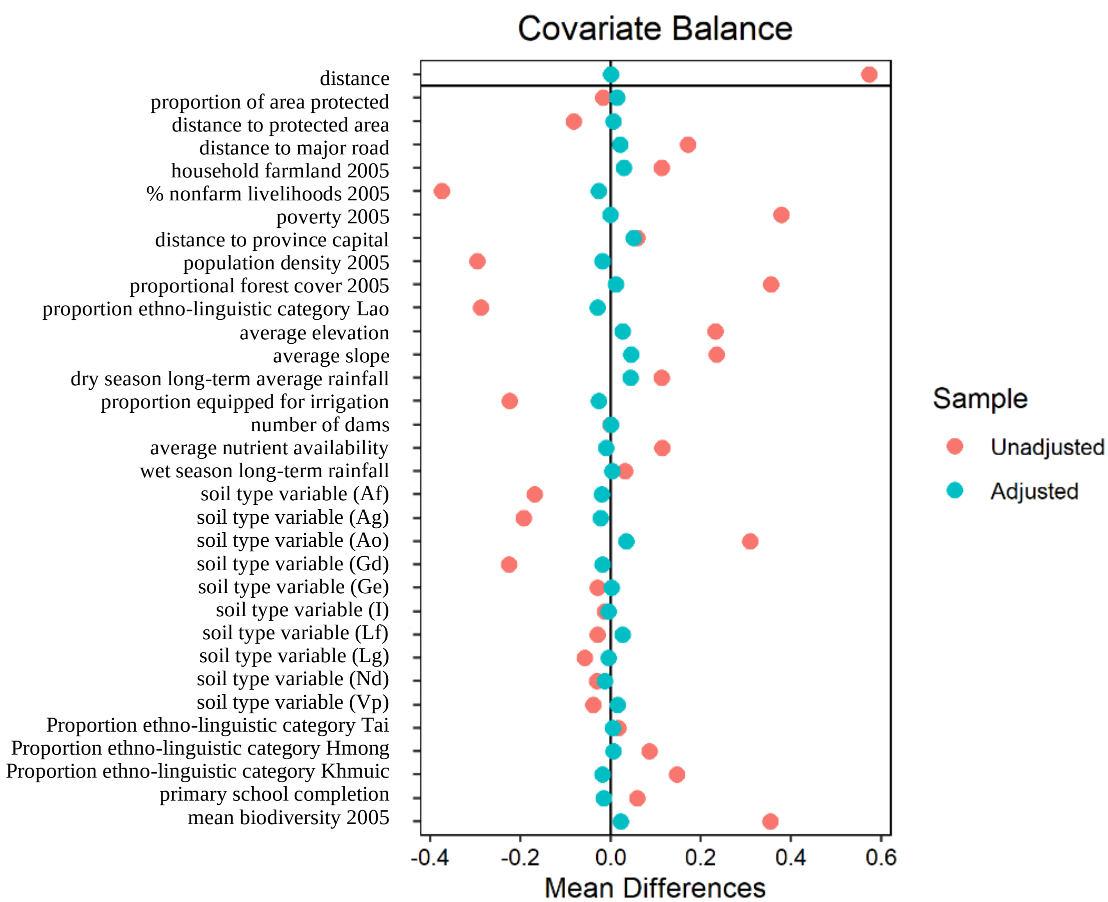


Figure 3.4: Love plot for all SDP projects PS model displaying the absolute standardized mean difference for all covariates.

3.2.3 SDP impacts on poverty

Overall, the IPWR results indicate that SDPs reduced poverty (first row of Table 3.2; the log of the poverty percentage was used as the outcome). For all SDP projects together, the average poverty reduction amounts to a significant 4.2%. According to the 95% confidence interval, our data are compatible with a poverty reduction between 1.8% and 6.7%. For food security and livelihoods projects, we obtain an average poverty reduction of 3.7%, with a confidence interval ranging from a 0.6% to 6.9% reduction in poverty. The results for sustainable forest management projects are less conclusive; our data are compatible with a poverty reduction up to 10.4% but also with a slight poverty increase up to 0.3%. Finally, for resource management projects, the confidence interval (-6.1% to 7.1%) does not rule out the absence of on SDP effect (which might also be due to the reduced sample size for this type of SDP). Spatial autocorrelation is relatively low for all poverty models. Additionally, our e-values analysis indicates some sensitivity to unobserved confounding (see Appendix A2.6).

Table 3.2: Regression outcomes estimating the average treatment effect on the treated (ATT), bootstrapped (R=5000) confidence intervals.

Outcome		All projects (n=1,452)	Food security and livelihoods projects (n= 829)	Sustainable forest management (n= 166)	Sustainable resource management (n= 375)
Log Poverty 2015 (%)	Treatment Effect	-4.2*	-3.7*	-5.0	0.5
		[-6.7, -1.8]	[-6.9, -0. 6]	[-10.4, 0.3]	[-6.1, 7.1]
Proportional forest cover (2015) †		0.9*	0.9*	0.0	1.8*
		[0.4, 1.3]	[0.4, 1.4]	[-0.9, 0.9]	[0.8, 2.8]

* $p < 0.05$ †outcome in percentage points

3.2.4 SDP impacts on forest cover

Three of the SDP models show positive and significant effects on 2015 forest cover – the models for all SDPs, food security and livelihoods, and resource management projects (second row of Table 3.2). For all SDP projects together, the average forest cover maintenance effect is .9 percentage points (pp). According to the 95% confidence interval, our data are compatible with a forest cover maintenance effect between .4 pp and 1.3 pp. For food security and livelihoods projects, we obtain an average forest cover maintenance effect of .9pp, with the confidence interval ranging from .4 pp to 1.4 pp. Similarly to the poverty models, the results for sustainable forest management projects are less conclusive; our data are compatible with a decrease in forest cover maintenance up to -.9pp but also with an increase in forest cover maintenance of .9pp. Resource management projects have the highest positive effect, resulting in a significant 1.8pp increase in proportional forest cover maintenance. The confidence interval suggests an ATT of between .8 to 2.8pp. Overall, spatial autocorrelation is relatively low for all forest cover models (Appendix A2.6.2). Additionally, our e-values analysis indicates some sensitivity to unobserved confounding (Appendix A2.6.1).

3.3 Discussion

3.3.1 Poverty

Findings from our IPWR analysis corroborate previous empirical analyses of SDPs and poverty, particularly those that examine food security and livelihood projects. Both the all-SDPs and the food security and livelihoods projects models indicate that SDPs reduced poverty. The magnitude of poverty alleviation is comparable to other recent studies on sustainability initiatives, specifically community-based forest management (Oldekop et al. 2019). However, our results show that these effects vary by program type. We find no significant impacts on poverty in our models of sustainable forest and resource management SDPs.

There are several potential explanations for the lack of poverty impacts for forest management SDPs. First, previous research in Laos has shown that smallholder forest plantations are not equally beneficial for all participants. Households with greater access to land and off-farm income sources have more success, and knowledge of production and marketing can be a challenge (Newby et al. 2012, van der Meer Simo 2020). Land tenure insecurity, market conditions (e.g. trader arrangements and fees), and limited producer organizations have also been identified as challenges for producers (Greijmans, Oudomvilay, and Banzon 2007, World Bank 2018). Another potential challenge facing forest SDPs is insufficient incentives for participants (World Bank 2018, World Wide Fund For Nature 2013). SDP documents suggest that logging restrictions in Laos during the study period fostered uncertainty in timber-focused SDPs, negatively impacting participant benefits (World Bank 2018).

For resource management SDPs, our finding of no significant effect is not surprising. Previous research has suggested varying effects of these interventions, ranging from increased poverty via unintended consequences or diminished access to natural resources (Barrett, Travis, and Dasgupta 2011, Martin, Myers, and Dawson 2018, Andersson et al. 2016) to improved human well-being due to resource conservation (Naidoo et al. 2019, Ferraro, Hanauer, and Sims 2011, Andam et al. 2010). In this case, though, the integrated conservation and development paradigm and participatory approaches utilized in Laos' SDPs may have limited negative poverty impacts. Arguably, habitat conservation projects may also provide co-benefits vis-à-vis the relationship between reduced deforestation and NTFP access, which we do not explicitly capture in this analysis (Choocharoen et al. 2013, Rigg 2006).

3.3.2 Forest cover

Our forest cover results comport with previous studies identifying positive environmental effects of SDPs (Honey-Rosés, Baylis, and Ramírez 2011, Ferraro and Simorangkir 2020). Three of our four models suggest that SDPs positively impacted forest cover during the study period, with resource management projects having the largest positive effect. However, sustainable forest management projects, when assessed separately, had no effect.

The positive impact of food security and livelihoods projects on forest cover suggests that efforts to provide alternative livelihoods options and dis-incentivize shifting cultivation were effective (see Asian Development Bank 2014). Notably, we find that food security and livelihoods projects achieved roughly half of the positive forest cover impact of resource management projects. Previous studies have suggested positive forest cover impacts of food and livelihoods SDPs due to, for example, provision of off-farm employment opportunities, increased agricultural productivity, and increased access to market-sourced goods (reducing the need for forest-sourced goods; Shively and Pagiola 2004, Paul J. Ferraro and Simorangkir 2020).

There are several potential explanations for our finding that forest management SDPs had no effect on forest cover. First, forest management SDPs may not have addressed key drivers of forest loss; large-scale challenges like illegal logging, concessions development, and local resource extraction may have prevented SDPs from maintaining forest cover (Singh 2020, Dwyer 2017, Hyakumura 2010). Second, lack of knowledge regarding production forest management and challenges regarding plantation registration for smallholders may have further hampered success (Pachas et al. 2019). Third, timber harvest in areas with high amounts of productive forestry may cyclically affect forest cover. While this is not necessarily associated with permanent deforestation, it could mask positive forest cover effects in our results. Finally, our analysis may have been limited by the smaller sample size for forest management projects.

3.3.3 Policy relevance and future work

Given ongoing debate in the sustainability community regarding the selection and design of voluntary interventions, our results strengthen the evidence base for positive impacts of SDPs and their continued use by international agencies, governments, NGOs, and local communities. Additionally, our results provide specific effect estimates that can inform realistic target setting and strategy development by stakeholders (Pressey et al. 2021). However, SDP project documents themselves acknowledge that previous evaluations were limited by a lack of baseline and monitoring data (World Bank 2018). Therefore, multilateral agencies and major donors should strongly consider making counterfactual assessments mandatory reporting requirements for SDPs. Combining these analyses with budgetary data could also shed light on cost-effectiveness of various interventions.

The co-benefits of food security and livelihoods SDPs present a promising topic for future applied research (Ferraro and Simorangkir 2020). However, studies rigorously assessing the impact of these projects on both poverty and forest cover are rare (D. Roe et al. 2015, Shively and Pagiola 2004). Further research in the Greater Mekong could explore income, forest cover, and food security outcomes in mixed tree crop systems (Phimmavong et al. 2019). From an implementation perspective, pairing forest management interventions with food security and livelihoods-focused efforts – e.g. via collaborative investment in agroforestry – may also maximize potential benefits (Phimmavong et al. 2019). Our results suggest that supporting food security and livelihoods SDPs may be beneficial even if forest cover, rather than poverty reduction, is the primary outcome of interest.

This study clearly demonstrates the benefits of counterfactual evaluation in identifying sustainability outcomes and informing program improvements. Although our results for forest management projects are inconclusive, these SDPs are heavily prioritized in Laos (Smith et al. 2021). For this reason, and due to their importance for land-based climate change emissions reduction, there is an urgent need for counterfactual and field-based research to investigate forest management SDP outcomes. Finally, voluntary SDPs must also be underpinned by strong government policy and corporate responsibility to amplify positive outcomes. This includes provision of extension services, clear tenure arrangements strengthening smallholder land and resource rights, and regulation to limit negative impacts of large-scale concessions and illegal resource extraction (Pachas et al. 2019, Singh 2020, Hett et al. 2020).

3.4 Materials and methods

We evaluated the impacts of 72 SDPs in 1,452 villages in Laos from 2005-2015 (SDPs are outlined in Table A2.5). Since there are no official national village boundaries available for Laos, we aggregated forest change and poverty data based on a set of 2015 village-level land-use area polygons, following Epprecht et al. 2008, Messerli et al. 2008, and the 2015 Population and Housing Census (see Appendix A2.2).

The analysis included the following steps: (1) drawing the directed acyclic graphs (DAGs) from subject matter theory, (2) articulating the project typology, (3) outcome variable calculation, (4) covariate selection and geospatial processing (e.g. calculating area weighted mean values; see Appendix A.2.2), (5) doubly robust IPWR estimation of the average treatment effect on the treated (ATT) with IPTWs calculated from propensity scores estimated via logistic regression, and (6) sensitivity analyses to probe the results' robustness to potentially unobserved confounding. First, DAGs were created based on subject matter theory regarding drivers affecting SDPs, poverty, and forest cover change, as well as a literature review of publicly available SDP project documents (see Appendix A2.1). This involved identifying factors influencing selection into SDPs (e.g. baseline poverty, baseline forest cover, biodiversity) as well as variables affecting sustainability outcomes (infrastructure development, government policy). We then generated an SDP project typology based on content analysis of project documents (Appendix A2.3). We defined four thematic groupings to reflect the SDPs primary activities and sustainability goals: (1) capacity development, (2) food security and livelihoods, (3) sustainable forest management, and (4) sustainable resource management. The project typology illustrated SDPs' different theories of change, ensuring that all projects included were designed to address forest cover and/or poverty outcomes. It also enabled us to assess the ATT for the different project types separately (apart from capacity development due to its low sample size), identifying potential effect heterogeneities.

The outcome variables include: (1) 2015 poverty headcount in percent, and (2) 2015 percent forest cover. Poverty data was derived from the national Laos 2015 Population and Housing Census, downloaded through the Lao DECIDE info platform (Lao DECIDE Info Project 2020). The census provides poverty percentage estimates at village level, which we log transformed prior to the IPWR analysis because of the dataset's right-skewed distribution. Moreover, poverty levels across villages more likely change proportional to the poverty base level, that is, in percentage terms rather than in absolute terms (percentage points). 2015 data included 8,404 villages in Laos, and excluded data from Xaisomboun Province, which was established in 2013. Forest cover data was derived from the Landsat data archive following methodology developed by Potapov et al. 2019 and Turubanova et al. 2018. We divided forest cover by land area to calculate percent cover (see Appendix A2.2.1).

Next, we selected covariates based on the DAGs and the adjustment criterion to remove confounding bias (Shpitser and VanderWeele 2011). We used the R package dagitty to conduct the graphical analysis and determine the adjustment set (Textor et al. 2016, see A2.1). The adjustment set consisted of 32 covariates for each model and included baseline measures of the outcome variables of interest and variables influencing selection into SDPs. The covariates related to broader themes including villages' socioeconomic status, sociopolitical context, market access, agricultural productivity, infrastructure, and environmental/biophysical conditions. Several of these variables are proxies for more nuanced factors included in the DAGs. For example, following previous studies, we used rainfall, soil type, and nutrient availability datasets to reflect agricultural productivity (Santika et al. 2019). Additionally, no post-intervention variables were added because they could have been influenced by SDP treatment (Schafer and Kang 2008). Table A2.4 in Appendix A2.2 shows the list of covariates, data sources, and processing approaches. All variables were aggregated to 2015 village polygons following the process described in A2.2.

The causal estimand of interest is the ATT, that is, the average treatment effect for the SDP villages. Using potential outcomes notation of the Rubin causal model (Rubin 1974, Imbens and Rubin 2015), the ATT is defined as the expected difference in potential treatment outcome (Y1) and control outcome (Y0), where the expectation is taken over all SDP villages (Z = 1):

$$ATT = \mathrm{E}(Y^1 - Y^0|Z = 1)$$

The ATT is causally identified if the adjustment set of covariates meets the adjustment criterion, i.e., blocks all confounding associations and meets the positivity assumption (Shpitser and VanderWeele 2011). The causal assumptions of this approach are explained and defended in Appendix A2.4.

Before estimating the ATT via IPWR, we estimated the propensity scores and IPTWs using the optimal full matching procedure of the *MatchIt* package in R (Ho et al. 2011, Hansen and Klopfer 2006). We estimated the propensity score via logistic regression with all covariates of the adjustment set included as predictors, created an optimal full matching stratification based on the propensity score, and then computed stratum-specific ATT weights for all control villages (all treatment villages received a weight of 1). We also assessed whether the ATT weights sufficiently balanced the covariate distributions of the treatment and control group using standardized mean differences and variance ratios, and visualized the achieved balance via love plots from the *cobalt* package (Steiner and Cook 2013, Greifer 2022). We used a caliper width of .2 standard deviations of the logit of the propensity score and trimmed any weights greater than

15 to 15 (Austin 2011, Lee, Lessler, and Stuart 2011). To facilitate the ATT estimation for the SDP types, we estimated separate propensity scores and weighted regressions for each type separately. This resulted in 8 final models assessing poverty and forest cover outcomes for (1) all SDPs together, (2) food security and livelihoods projects, (3) sustainable forest management projects, and (4) sustainable resource management projects (see Appendix A2.5.1).

We then used the doubly robust IPWR analysis to estimate the ATT. IPWR combines IPTWs with an additional regression adjustment and thus provides us two chances to obtain a consistent estimate: if the functional form of either the PS model or the regression model for the outcome is correctly specified, a consistent estimate will result (Lunceford and Davidian 2004). IPWR is implemented as weighted least squares regression with IPTW weights (Imbens and Wooldridge 2009, Steiner and Cook 2013):

$$Y = \alpha + \tau Z + X'\beta + \varepsilon$$

where Z is the treatment indicator, τ the ATT, X the vector of mean centered covariates of the adjustment set, and β the corresponding coefficient vector. Due to multicollinearity we dropped 11 covariates with a variance inflation factor greater than 10 from the adjustment set for the regression (but they remained in the PS model). We utilized the *sandwich* package in R to bootstrap 95% confidence intervals for the effect estimates. These were directly obtained from the bootstrap distributions based on 5,000 replications (Zeileis, Köll, and Graham 2020, Zeileis 2004). Model code will be available in the Supplemental Materials upon publication.

By determining the adjustment set of covariates from the DAGs and using a doubly robust IPWR estimator, we aimed to minimize any confounding bias to the extent possible. Nonetheless, some confounding bias could remain due to unobserved or unreliably measured confounders (see A2.1, A2.4). However, based on subject matter theory and the number and diversity of available covariates, we argue that any remaining bias should be negligibly small because (1) confounding constructs have been measured with high reliability; that is, each construct, even if not perfectly measured, is covered by at least one close proxy and several more distant proxy variables, and (2) for the unobserved confounders government policy and other sustainability projects, several of the observed covariates are presumably highly associated and thus remove most of their bias. Following best practice, we also implemented a sensitivity analysis to probe the effect estimates' robustness to unobserved confounding using e-values calculations (VanderWeele and Ding 2017, Mathur et al. 2021, see A2.6.1). Finally, we tested for spatial autocorrelation by calculating Moran's I of the post-matching regression residuals, following Oldekop et al. 2019 and Schleicher et al. 2020 (see A2.6.2).

3.5 Bibliography

Ashraf, Jawaid, Rajiv Pandey, and Wil de Jong. 2017. “Assessment of Bio-Physical, Social and Economic Drivers for Forest Transition in Asia-Pacific Region.” *Forest Policy and Economics*, Forest transition in Asia, 76 (March): 35–44. https://doi.org/10.1016/j.forpol.2016.07.008.

Asian Development Bank. 2004. “Northern Community-Managed Irrigation Sector Project: Report and Recommendation of the President to the Board of Directors on a Proposed Loan.” RRP: LAO 34188. Asian Development Bank. https://www.adb.org/sites/default/files/project-document/69972/rrp-lao-34188.pdf.

———. 2006. “Road Network Map of Lao PDR.”

———. 2007. “Northern Region Sustainable Livelihoods through Livestock Development Project.” 35297. Project Administration Memorandum. Asian Development Bank. https://www.adb.org/sites/default/files/project-document/66425/35297-lao-pam.pdf.

———. 2008. “Alternative Livelihood for Upland Ethnic Groups in Houaphanh Province (Financed by the Japan Fund for Poverty Reduction).” 41096. Grant Assistance Report. Asian Development Bank. https://www.adb.org/sites/default/files/project-document/65334/41096-lao-gar.pdf.

Bader, Christoph, Sabin Bieri, Urs Wiesmann, and Andreas Heinimann. 2016. “Differences Between Monetary and Multidimensional Poverty in the Lao PDR: Implications for Targeting of Poverty Reduction Policies and Interventions.” *Poverty & Public Policy* 8 (2): 171–97. https://doi.org/10.1002/pop4.140.

Baird, Ian G., and Bruce Shoemaker. 2007. “Unsettling Experiences: Internal Resettlement and International Aid Agencies in Laos.” *Development and Change* 38 (5): 865–88. https://doi.org/10.1111/j.1467-7660.2007.00437.x.

Bivand, Roger, Micah Altman, Luc Anselin, Renato Assunção, Olaf Berke, Guillaume Blanchet, Marilia Carvalho, et al. 2022. *Spdep: Spatial Dependence: Weighting Schemes, Statistics* (version 1.2-4). https://CRAN.R-project.org/package=spdep.

Bouahom, Bounthong, Linkham Douangsavanh, and Jonathan Rigg. 2004. “Building Sustainable Livelihoods in Laos: Untangling Farm from Non-Farm, Progress from Distress.” *Geoforum*, Themed section on The Spaces of Critical Geography, 35 (5): 607–19. https://doi.org/10.1016/j.geoforum.2004.02.002.

Broegaard, Rikke Brandt, Thoumthone Vongvisouk, and Ole Mertz. 2017. “Contradictory Land Use Plans and Policies in Laos: Tenure Security and the Threat of Exclusion.” *World Development* 89 (January): 170–83. https://doi.org/10.1016/j.worlddev.2016.08.008.

Elwert, Felix. 2013. “Graphical Causal Models.” In *Handbook of Causal Analysis for Social Research*, edited by Stephen L. Morgan, 245–73. Handbooks of Sociology and Social Research. Dordrecht: Springer Netherlands. https://doi.org/10.1007/978-94-007-6094-3_13.

Epprecht, Michael, Nicholas Bosoni, Albrecht Ehrensperger, Hiroko Nagasawa, Juliet Lu, David Studer, Patrizia Vollmar, Rasso Bernhard, and Vilaysook Sisoulath. 2018. *Socio-Economic Atlas of the Lao PDR. Patterns and Trends from 2005 to 2015*. Bern, Switzerland and Vientiane, Lao PDR: Centre for Development and Environment, University of Bern, and Lao Statistics Bureau, Lao PDR, with Bern Open Publishing. https://doi.org/10.7892/boris.120671.

Epprecht, Michael, Nicolas Minot, Reno Dewina, Peter Messerli, and Andreas Heinimann. 2008. "The Geography of Poverty and Inequality in the Lao PDR." Application/pdf. https://doi.org/10.7892/BORIS.34330.

Ferraro, Paul J., and Rhita Simorangkir. 2020. "Conditional Cash Transfers to Alleviate Poverty Also Reduced Deforestation in Indonesia." *Science Advances* 6 (24): eaaz1298. https://doi.org/10.1126/sciadv.aaz1298.

Fischer, G., F. Nachtergaele, S. Prieler, H. T. van Velthuizen, L. Verelst, and D. Wiberg. 2008. "Global Agro-Ecological Zones Assessment for Agriculture. Dataset." Laxenburg, Austria / Rome, Italy: IIASA and FAO. https://webarchive.iiasa.ac.at/Research/LUC/External-World-soil-database/HTML/SoilQualityData.html?sb=11.

Food and Agriculture Organization of the United Nations. 2007. "FAO Digital Soil Map of the World (Version 3.6). Dataset." 2007. https://data.apps.fao.org/map/catalog/srv/eng/catalog.search#/metadata/446ed430-8383-11db-b9b2-000d939bc5d8.

Global Environment Facility. 2004. "Integrated Ecosystem and Wildlife Management in Bolikhamxay Province - Project Proposal." MEDIUM-SIZED PROJECT PROPOSAL. Global Environment Facility. https://documents1.worldbank.org/curated/en/395381468772454991/pdf/316500LaoPDR0MSP0Brief012113104.pdf.

Greifer, Noah. 2022. *Cobalt: Covariate Balance Tables and Plots* (version 4.3.2). https://CRAN.R-project.org/package=cobalt.

Harris, I., P.D. Jones, T.J. Osborn, and D.H. Lister. 2014. "Updated High-Resolution Grids of Monthly Climatic Observations - the CRU TS3.10 Dataset: UPDATED HIGH-RESOLUTION GRIDS OF MONTHLY CLIMATIC OBSERVATIONS." *International Journal of Climatology* 34 (3): 623–42. https://doi.org/10.1002/joc.3711.

Heilmayr, Robert, Kimberly M. Carlson, and Jason Jon Benedict. 2020. "Deforestation Spillovers from Oil Palm Sustainability Certification." *Environmental Research Letters* 15 (7): 075002. https://doi.org/10.1088/1748-9326/ab7f0c.

Ho, Daniel, Kosuke Imai, Gary King, and Elizabeth A. Stuart. 2011. "MatchIt: Nonparametric Preprocessing for Parametric Causal Inference." *Journal of Statistical Software* 42 (1): 1–28. https://doi.org/10.18637/jss.v042.i08.

ICEM. 2003. "Lao People's Democratic Republic: National Report on Protected Areas and Development." Review of Protected Areas and Development in the Four Countries of the Lower Mekong River Region. Indooroopilly, Queensland, Australia: International Centre for Environmental Management (ICEM). https://icem.com.au/documents/biodiversity/pad/lao_pdr_nr.pdf.

Imbens, Guido W. 2003. "Nonparametric Estimation of Average Treatment Effects under Exogeneity: A Review." t0294. National Bureau of Economic Research. https://doi.org/10.3386/t0294.

Jones, Sam, and Peter Gibbon. 2011. "Developing Agricultural Markets in Sub-Saharan Africa: Organic Cocoa in Rural Uganda." *The Journal of Development Studies* 47 (10): 1595–1618. https://doi.org/10.1080/00220388.2011.579107.

Junquera, Victoria, Patrick Meyfroidt, Zhanli Sun, Phokham Latthachack, and Adrienne Grêt-Regamey. 2020. "From Global Drivers to Local Land-Use Change: Understanding the

Northern Laos Rubber Boom." *Environmental Science & Policy* 109 (July): 103–15. https://doi.org/10.1016/j.envsci.2020.04.013.
Knowledge for Development. 2022. "K4D Map Explorer." 2022. https://apps.k4d.la/explorer/.
Lao DECIDE Info Project. 2020. "Lao DECIDE Info - Start." Lao DECIDE Info. 2020. http://www.decide.la/en/.
Manorom, Kanokwan, Ian G. Baird, and Bruce Shoemaker. 2017. "The World Bank, Hydropower-Based Poverty Alleviation and Indigenous Peoples: On-the-Ground Realities in the Xe Bang Fai River Basin of Laos." *Forum for Development Studies* 44 (2): 275–300. https://doi.org/10.1080/08039410.2016.1273850.
Martin, Sarah M., and Kai Lorenzen. 2016. "Livelihood Diversification in Rural Laos." *World Development* 83 (July): 231–43. https://doi.org/10.1016/j.worlddev.2016.01.018.
Mathur, Maya B., Louisa H. Smith, Peng Ding, and Tyler J. VanderWeele. 2021. *EValue: Sensitivity Analyses for Unmeasured Confounding and Other Biases in Observational Studies and Meta-Analyses* (version 4.1.3). https://CRAN.R-project.org/package=EValue.
Mekong Region Futures Institute. 2021. "Dataset on the Dams of the Greater Mekong." Bangkok: Mekong Region Futures Institute.
Messerli, Peter, Andreas Heinimann, Michael Epprecht, Phonesaly Souksavath, Thiraka Chanthalanouvong, and Nicholas Minot. 2008. *Socio-Economic Atlas of the Lao PDR. An Analysis Based on the 2005 Population and Housing Census.* Bern: Geographica Bernensia. http://www.laoatlas.net/links/PDF%20download.html.
Morgan, Stephen L., and Christopher Winship. 2014. *Counterfactuals and Causal Inference: Methods and Principles for Social Research*. 2nd ed. Analytical Methods for Social Research. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9781107587991.
Nanhthavong, Vong, Michael Epprecht, Cornelia Hett, Julie G. Zaehringer, and Peter Messerli. 2020. "Poverty Trends in Villages Affected by Land-Based Investments in Rural Laos." *Applied Geography* 124 (November): 102298. https://doi.org/10.1016/j.apgeog.2020.102298.
Newbold, Tim, Lawrence Hudson, Andy Arnell, Sara Contu, Adriana De Palma, Simon Ferrier, Samantha L. L. Hill, et al. 2016. "Global Map of the Biodiversity Intactness Index, from Newbold et al. (2016) Science," July. https://doi.org/10.5519/0009936.
Oldekop, Johan A., Katharine R. E. Sims, Birendra K. Karna, Mark J. Whittingham, and Arun Agrawal. 2019. "Reductions in Deforestation and Poverty from Decentralized Forest Management in Nepal." *Nature Sustainability* 2 (5): 421–28. https://doi.org/10.1038/s41893-019-0277-3.
Open Development Mekong. 2010. "Laos Protected Areas and Heritage Sites- OD Mekong Datahub." 2010. https://data.opendevelopmentmekong.net/dataset/laos-protected-areas-and-heritage-sites/resource/c4333eba-f031-45cd-bd58-167709847369?inner_span=True.
Pearl, Judea. 1993. "[Bayesian Analysis in Expert Systems]: Comment: Graphical Models, Causality and Intervention." *Statistical Science* 8 (3): 266–69.
———. 2009. *Causality*. 2nd ed. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511803161.
Pebesma, Edzer, Roger Bivand, Etienne Racine, Michael Sumner, Ian Cook, Tim Keitt, Robin Lovelace, et al. 2022. *Sf: Simple Features for R* (version 1.0-7). https://CRAN.R-project.org/package=sf.

Pfaff, Alexander, and Juan Robalino. 2017. “Spillovers from Conservation Programs.” *Annual Review of Resource Economics* 9 (1): 299–315.
Phompila, Chittana, Megan Lewis, Bertram Ostendorf, and Kenneth Clarke. 2017. “Forest Cover Changes in Lao Tropical Forests: Physical and Socio-Economic Factors Are the Most Important Drivers.” *Land* 6 (2): 23. https://doi.org/10.3390/land6020023.
Potapov, Peter, A. Tyukavina, S. Turubanova, Y. Talero, A. Hernandez-Serna, M. C. Hansen, D. Saah, et al. 2019. “Annual Continuous Fields of Woody Vegetation Structure in the Lower Mekong Region from 2000-2017 Landsat Time-Series.” *Remote Sensing of Environment* 232 (October): 111278. https://doi.org/10.1016/j.rse.2019.111278.
R Core Team. 2020. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. https://www.R-project.org.
Rigg, J. D. 2006. “Forests, Marketization, Livelihoods and the Poor in the Lao PDR.” *Land Degradation & Development* 17 (2): 123–33. https://doi.org/10.1002/ldr.719.
Robinson, Brian E., Margaret B. Holland, and Lisa Naughton-Treves. 2014. “Does Secure Land Tenure Save Forests? A Meta-Analysis of the Relationship between Land Tenure and Tropical Deforestation.” *Global Environmental Change* 29 (November): 281–93. https://doi.org/10.1016/j.gloenvcha.2013.05.012.
Rohrer, Julia M. 2018. “Thinking Clearly About Correlations and Causation: Graphical Causal Models for Observational Data.” *Advances in Methods and Practices in Psychological Science* 1 (1): 27–42. https://doi.org/10.1177/2515245917745629.
Rubin, Donald B. 1978. “Bayesian Inference for Causal Effects: The Role of Randomization.” *The Annals of Statistics* 6 (1): 34–58. https://doi.org/10.1214/aos/1176344064.
———. 2001. “Using Propensity Scores to Help Design Observational Studies: Application to the Tobacco Litigation.” *Health Services and Outcomes Research Methodology* 2 (3): 169–88. https://doi.org/10.1023/A:1020363010465.
Schafer, Joseph L., and Joseph Kang. 2008. “Average Causal Effects from Nonrandomized Studies: A Practical Guide and Simulated Example.” *Psychological Methods* 13 (4): 279–313. https://doi.org/10.1037/a0014268.
Siebert, Stefan, Verena Henrich, Karen Frenken, and Jacob Burke. 2013. “Global Map of Irrigation Areas Version 5. Dataset.” Bonn, Germany / Rome, Italy: Rheinische Friedrich-Wilhelms-University / Food and Agriculture Organization of the United Nations. https://www.fao.org/aquastat/en/geospatial-information/global-maps-irrigated-areas/latest-version.
Simo, Alex Van Der Meer, Peter Kanowski, and Keith Barney. 2019. “Revealing Environmental Income in Rural Livelihoods: Evidence from Four Villages in Lao PDR.” *Forests, Trees and Livelihoods* 28 (1): 16–33. https://doi.org/10.1080/14728028.2018.1552540.
Smith, Hilary, Juliet Lu, Phuc Xuan To, Soytavanh Mienmany, and Khonethong Soukphaxay. 2020. “Rubber Plantation Value Chains in Laos: Opportunities and Constraints in Policy, Legality and Wood Processing.” Forest Trends. https://www.forest-trends.org/wp-content/uploads/2020/07/Rubber-Plantation-Value-Chains-in-Laos.pdf.
Stanley, Scott Alexander, Richard Hackman, Jeremy Ferrand, and Forest Carbon Consultants Inc. 2013. “REDD Feasibility Study for Xe Sap, Saola, and Saola Quang Nam NPAs in Lao PDR and Vietnam.” World Wide Fund for Nature, Car-Bi Project.
Steiner, Peter M., and David Cook. 2013. “Matching and Propensity Scores.” In *The Oxford Handbook of Quantitative Methods*. Oxford University Press.

Stuart, Elizabeth A. 2010. “Matching Methods for Causal Inference: A Review and a Look Forward.” *Statistical Science: A Review Journal of the Institute of Mathematical Statistics* 25 (1): 1–21. https://doi.org/10.1214/09-STS313.

Textor, Johannes. 2015. “Drawing and Analyzing Causal DAGs with DAGitty.” http://www.dagitty.net/manual-2.x.pdf.

Textor, Johannes, Benito van der Zander, Mark S. Gilthorpe, Maciej Liskiewicz, and George Th Ellison. 2016. “Robust Causal Inference Using Directed Acyclic Graphs: The R Package ‘Dagitty.’” *International Journal of Epidemiology* 45 (6): 1887–94. https://doi.org/10.1093/ije/dyw341.

Thomas, Ian Lloyd. 2015. “USAID Lowering Emissions in Asia’s Forests (USAID LEAF): Drivers of Forest Change in the Greater Mekong Subregion. Laos Country Report.” USAID LEAF. https://www.uncclearn.org/wp-content/uploads/library/fao13102015_4.pdf.

Tong, Pei Sin. 2009. “LAO PEOPLE’S DEMOCRATIC REPUBLIC FORESTRY OUTLOOK STUDY.” Working Paper No. APFSOS II/WP/2009/17. Working Paper Series. Bangkok: Food And Agriculture Organization Of The United Nations.

Traldi, Rebecca, Julie A. Silva, Peter Potapov, Alexandra Tyukavina, Michael Epprecht, Meredith L. Gore, and Chittana Phompila. 2022. “Cultivating Inequality? Regional Rubber Dynamics and Implications for Voluntary Sustainability Programs in Lao PDR.” *Under Review*.

Turubanova, Svetlana, Peter V. Potapov, Alexandra Tyukavina, and Matthew C. Hansen. 2018. “Ongoing Primary Forest Loss in Brazil, Democratic Republic of the Congo, and Indonesia.” *Environmental Research Letters* 13 (7): 074028. https://doi.org/10.1088/1748-9326/aacd1c.

UNEP-WCMC. 2022. “Protected Planet.” Protected Planet. 2022. https://www.protectedplanet.net/en.

VanderWeele, Tyler J., and Peng Ding. 2017. “Sensitivity Analysis in Observational Research: Introducing the E-Value.” *Annals of Internal Medicine* 167 (4): 268. https://doi.org/10.7326/M16-2607.

Vongvisouk, Thoumthone, Rikke Brandt Broegaard, Ole Mertz, and Sithong Thongmanivong. 2016. “Rush for Cash Crops and Forest Protection: Neither Land Sparing nor Land Sharing.” *Land Use Policy* 55 (September): 182–92. https://doi.org/10.1016/j.landusepol.2016.04.001.

Wade, Len, Tamara Jackson, Jenny Locker, and Sarah Boothey. 2015. “Developing Improved Farming and Marketing Systems in Rainfed Regions of Southern Lao PDR - Final Report.” CSE/2009/004. Australian Centre for International Agricultural Research. https://aciar.gov.au/publication/technical-publications/developing-improved-farming-and-marketing-systems-rainfed-regions-southern-lao-pdr.

Wieder, W. 2014. “Regridded Harmonized World Soil Database v1.2,” 59.234908 MB. https://doi.org/10.3334/ORNLDAAC/1247.

World Bank. 2018. “Lao People’s Democratic Republic- Sustainable Forestry for Rural Development Project.” 125929. Independent Evaluation Group, Project Performance Assessment Report. Washington DC: The World Bank. https://documents1.worldbank.org/curated/fr/239161528225963013/pdf/125929-PPAR-P163409-P064886-PUBLIC.pdf.

World Food Programme, and Open Street Map. 2018. "Laos Road Network (Main Roads). Dataset." WFP GeoNode. 2018. https://geonode.wfp.org/layers/ogcserver.gis.wfp.org%3Ageonode%3Alao_trs_roads_osm.

World Wide Fund For Nature. 2013. "Results from the CarBi REDD Feasibility Study | WWF." World Wide Fund For Nature. 2013. https://wwf.panda.org/wwf_news/?208939/Results-from-the-CarBi-REDD-Feasibility-Study.

Chapter 4: Cultivating Inequality? Regional rubber dynamics and implications for voluntary sustainability programs in Lao PDR[19]

4.1 Introduction

Inequality has long been a significant social concern and manifests in different ways, including along economic, ethnic, and gender lines. While not representative of all forms of inequality, economic inequality – essentially, differences in the income shares captured by different groups or individuals – often serves as an entry point to better understand uneven development (Osberg 2015, Silber 2012, Chancel et al. 2022). Inequality is an important phenomenon due to its potential to both stem from and drive conflict and other negative social consequences. Additionally, it is often connected to human and sustainable development challenges, including land use change, environmental deterioration, and poverty (Milanovic 2016, Baek and Gweisah 2013, Scruggs 1998, Basu 2006, Ravallion 2005). Consequently, inequality is integrated into diverse global sustainable development frameworks, and designated as one of the 17 United Nations Sustainable Development Goals (SDG 10: Reduced Inequalities).

Although less studied in the development literature, economic polarization is another important aspect of income distribution with broader social implications. While economic inequality focuses on the spread of an income distribution, polarization is designed to investigate clustering around local means (Zhang and Kanbur 2001, Silva et al. 2015). Polarization measures

---

[19] Note that this chapter is under review in *World Development* (submitted January 2022). Co-authors of this chapter include Dr. Julie A. Silva, Dr. Peter Potapov, Dr. Alexandra Tyukavina, Dr. Michael Epprecht, Dr. Meredith L. Gore, and Dr. Chittana Phompila.

can indicate distributional dynamics linked to social conflict that are not revealed by inequality indices, such as a disappearing middle class (Zhang and Kanbur 2001, Ricci and Scicchitano 2021). Thus, the different foci of the two metrics mean the indices can move in opposite directions. For example, economic inequality may decline in response to economic shocks that disproportionately affect wealthier households, but polarization could still rise if incomes become more tightly clustered at opposite ends of the existing distribution.

Rubber (*hevea brasilensis*) production – specifically, the Southeast Asian rubber boom from roughly 2000-2011 - provides a backdrop of rapid land-use change against which to examine trends in rural economic inequality. Spurred by market forces, policy, and agro-technological development, the rubber boom brought economic growth but also serious sustainability consequences, including those related to deforestation and local access to land (Ahrends et al. 2015, Baird 2010). For this reason, the crop has received increasing attention from the sustainability community, evidenced by new private sector sustainable sourcing commitments, the development of industry sustainability standards, and other activities (GPSNR 2020, IRSG 2014, Michelin 2021, Pirelli 2021, FSC 2021). Echoing broader trends in the sustainable agriculture and forestry realms, many of these initiatives are voluntary multi-stakeholder efforts, driven by industry and NGO collaboration, rather than government regulation (Willer et al. 2019).

Although there has been significant research assessing rubber production in the Mekong region, including its economic impacts, there has been less focus specifically on economic inequality and polarization outcomes in rubber producing areas. The limited research linking rubber production to economic inequality suggests that economic benefits from production may accrue unequally within communities and among stakeholders (e.g. Evans et al. 2011, Vongvisouk et al. 2014). Rubber is also grown via vastly different production systems. For example, in Lao PDR (Laos), rubber production areas range from smallholder plots of a few hectares to enormous monoculture plantations funded by foreign investments (Smith et al. 2020). Previous research suggests that inequality dynamics differ among these production system types, with large concessions in Southern Laos particularly likely to drive negative impacts (see Cramb et al. 2017, Fox and Castella 2013, Nanhthavong et al. 2022, *forthcoming*).

In the context of sustainable development, economic inequality and polarization may disincentivize participation in voluntary sustainability projects or limit these projects' ability to achieve sustainability goals (see Halpern et al. 2013, Law et al. 2018). Further, in the sustainability literature, there is frequent concern that sustainability interventions may reinforce pre-existing inequalities and power structures (Bray and Neilson 2017, Gardner et al. 2019, Pinto and McDermott 2013, McDermott and Schreckenberg 2009). However, the relationship of sustainability interventions to economic inequality is rarely assessed in practice (Kalonga, Kulindwa, and Mshale 2015).

Our research investigates changes in rural economic inequality and polarization in the context of rapid growth in rubber production and examines the implications of this for voluntary sustainability projects, using Laos as a case study. Specifically, we ask:

(1) To what extent was rubber production associated with changes in rural economic inequality and polarization in Laos from 2007/08-2012/13?

(2) What are the implications of these trends for voluntary rubber sustainability programs?

We focus on rural areas due to their importance for rubber production, as well as comparatively higher poverty rates when juxtaposed with urban areas (Warr, Rasphone, and Menon 2015). This also enables us to explore inequality among neighbors, rather than along rural and urban lines. By assessing the consequences of our findings for voluntary sustainability programs, we explore how these programs are situated in a context of changing economic inequality and polarization, and their options to implement design changes to maintain participation and mitigate unintended consequences. Since other studies have already explored policy frameworks and recommendations for sustainable rubber and forest plantations more generally (e.g. Smith et al. 2021, Smith et al. 2020), we focus primarily on voluntary programs.

Laos provides a compelling case study for investigating varied economic inequality impacts of a crop boom. First, increases in rubber area occurred within regionally differentiated production systems – specifically, smallholder, concessions, and contract farming - with the Northern region seeing more prominent smallholder production and the Southern region experiencing an influx of large-scale plantations (Smith et al. 2020). This diversity of production arrangements provides an opportunity to investigate how rubber's impact on rural economic inequality varies in different contexts. Additionally, since this analysis assesses a unique 5-year time period when rubber prices reached their peak and then steeply decreased (Vongvisouk and Dwyer 2016), it is uniquely positioned to assess inequality changes during a significant crop boom followed by a rapid price recession. Finally, rubber is known to drive socioeconomic challenges in Southeast Asia, including for workers and for local communities. Particularly on large plantations, problems have ranged from nonexistent or insufficient contracts; to low, piece-rate, and seasonal income for tappers; and lack of worker protection and health risks (Smith et al. 2020, Singh 2020, Kenney-Lazar 2012). Production has also been associated with disputes between local communities and investor companies, and allegations of land grabbing (Baird

2010, Schönweger et al. 2012). Analyzing rubber's impact on rural economic inequality and polarization provides an opportunity to explore these challenges.

We utilize several methods to assess changes in rural economic inequality and polarization within rubber and non-rubber producing areas in Laos, based on remotely sensed imagery and household expenditure surveys. Our methods include satellite-based rubber area estimation, and calculation of the Gini coefficient, Generalized Entropy indices, and the Duclos Esteban Ray (DER) polarization index. The remainder of the manuscript is structured as follows: First, we provide contextual background on inequality, rubber production, and relevant sustainability initiatives (Section 2). This is followed by a description of our methods and data sources (Section 3), and our results (Section 4). Finally, we explore the observed trends in rural economic inequality and polarization and offer insights related to rubber production in voluntary sustainability programs (Sections 5 and 6).

4.2 Case Study Background

4.2.1 Inequality in Laos

Over the last several decades, Laos has followed the general pattern of many other emerging economies, in that poverty overall has decreased while economic inequality has increased. At the national level, inequality in Laos as measured by the Gini coefficient increased from .31 to .36 between 1992 and 2013 (Pimhidzai et al. 2014, Warr, Rasphone, and Menon 2015). Studies of regional economic inequality indicate that between 1992 and 2013, the Gini coefficient increased by roughly 18% in the North, 16% in the South, and 6% in the Central region, reflecting higher inequality in the country (Warr, Rasphone, and Menon 2015). However, between 2007 and 2013, inequality increased only in the Southern region (Ibid). For a map of regions within Laos, see Figure 4.1.

Figure 4.1: Regions of Laos PDR (Laos). Data sources: Authors' elaboration based on data from Open Development Mekong (2021) and Esri Topographic map.

The 'New Economic Mechanism' policy, introduced in 1986 to open Laos up to trade and foreign investment, has contributed to sustained economic growth (Bader et al. 2017). During this period, economic inequality deepened along urban and rural divides – with urban areas generally seeing higher inequality – as well as along ethnic lines (Warr, Rasphone, and Menon 2015, Bader et al. 2017). Interestingly, some research has shown that higher-income rural areas in Laos (often located in lowland regions) have tended to have lower levels of economic inequality, while there have been pockets of high inequality in rural upland regions (e.g. in Bolikhamxai and Sekong), as well as in some low-income districts (Epprecht et al. 2008).

Like other countries in Southeast Asia, Laos has experienced significant transitions in its agricultural sector over the last few decades, with the government promoting more intensive and plantation-based agriculture – through mechanisms like the Land and Forest Allocation Policy (LFA) – amidst challenges with land tenure and poverty (Castella et al. 2013, Phimmavong et al. 2009, Baird 2014). Relatively few case studies evaluate the implications of these changes for inequality specifically in Laos, although many studies assess broader human well-being impacts (Junquera and Grêt-Regamey 2020, Friis et al. 2016). One study which explored transitions from shifting cultivation to rubber production through agent-based modeling found that after new rubber adoption, inequality remained relatively low for the six years from planting to initial tapping, but then increased due to high revenues for early adopters (Evans et al. 2011). This comports with research suggesting that large-scale tree plantation development has often increased inequality worldwide (Malkamäki et al. 2018). Additionally, previous research predicted that the tree plantation policy would increase inequality in Laos and provide some poverty reduction benefits (Phimmavong and Keenan 2020).

4.2.2 Rubber production in Laos

The growth of rubber production in Laos between 2000 and 2011 followed broader regional expansion in the Greater Mekong. Some factors driving expansion were exogenous to the Laotian context – for example, dramatic price increases for rubber, and heightened demand from the People's Republic of China (Kenney-Lazar et al. 2018). Previous research indicates that much of Laos' rubber expansion occurred between 2006 and 2011 (Hurni and Fox 2018). Local drivers of rubber expansion included land-use policies promoting cash crop production to replace alternative land uses, the influence of cross-border production and trade networks, and improvements in transportation infrastructure (Junquera et al. 2020, Lu 2017). Some rubber production occurred in Laos prior to the boom period, with the number of rubber growers increasing from 100 in 1999 to 49,000 in 2010/2011 (MAF 2014). According to previous research, rapid production increases occurred beginning in 2004 in the North, and 2006 in the South (Evans et al. 2011, Özdoğan, Baird, and Dwyer 2018, Epprecht et al. 2018). The top rubber-producing provinces in 2018 were Louangnamtha, Oudomxai, and Champasack, with 14%, 11% and 11% of planted area respectively (NAFRI 2018 in Smith et al. 2020, p.18). Previous research indicates that the rubber boom drove significant deforestation in Laos, as well as regionally. For example, Hurni and Fox (2018) found that 70% of rubber plantation development in mainland Southeast Asia resulted in deforestation.

Rubber trees take 6-7 years to begin producing latex, and only a small amount of latex is produced in years 8 to 10; production peaks at around 20 years (Junquera et al. 2020, Vongvisouk et al. 2014). By 2018, roughly 260,000 hectares of rubber had been planted in Laos, although there are some discrepancies in official figures, and recent studies based on remotely sensed data estimate higher rubber areas (Vongkhamor 2016 in Kenney-Lazar et al. 2018, Smith et al. 2020, Hurni and Fox 2018). Although rubber prices decreased steeply from 2011-2015 and have remained relatively low since then, global demand has remained strong due to the tire market and higher costs for synthetic alternatives driven by high oil prices (Vongvisouk and Dwyer 2016, Jin et al. 2021, Index Mundi 2021, Kenney-Lazar et al. 2018, Warren-Thomas, Dolman, and Edwards 2015). Expansion slowed in part due to market signals, as well as the concession moratoria implemented in 2007 and 2012 (Smith et al. 2020).

Rubber production systems in Laos can be broadly categorized into three groups: (1) smallholder systems, (2) plantations typically established through government concessions, and (3) contract farming arrangements between companies and farmers (Kenney-Lazar et al. 2018). Within all these arrangements, there is variability. For example, within contract farming, the design of the contract arrangement, a producer's previous experience growing rubber and the presence of local farmer groups may influence the negotiation and finalization of contracts (Fox et al. 2014). One source estimated that of rubber planted area in Laos, about 28% is smallholder agriculture, 28% is contract farming, and the remainder (about 44%) is concession-based (FSIS 2018 in Smith et al. 2020).

4.2.3 Regional rubber production systems: Drivers, structure, and impacts

Due to the relatively rapid expansion of rubber production in Laos, as well as its economic importance, there has been a substantial amount of research exploring the crop's environmental and social outcomes and drivers. This body of literature contextualizes the significant distinctions between rubber production systems in the country's Northern, Central, and Southern regions.

4.2.3.1 Rubber in Northern Laos

Rubber production in Northern Laos is characterized by smallholder systems (<3 ha), with rubber uptake influenced by knowledge-sharing, imitation, and social networks (Baird and Vue, 2017, Junquera et al. 2020, Douangsavanh, 2009). Some households in Louangnamtha began planting rubber in 1994, although the boom began in earnest later about a decade later

(Evans et al. 2011). According to the 2010/11 Agricultural Census, 15% of farmers in the North had planted rubber trees (WFP, 2013). Northern smallholder systems are a blend of private investment, government concessions, sharecropping, and contract farming (Baird and Vue, 2017). For example, Louangnamtha's rubber area is primarily managed by individual producers followed by concessions, while production in Oudomxai is almost entirely based on contract farming (Southavilay 2016 in Smith et al. 2020, p.21). Some research has suggested that smallholder production generates greater economic benefits as compared to other land uses in the North, and several studies argue that rubber production was strategically utilized by local communities to secure land titles (Nanhthavong et al. 2020, Fox et al. 2014, Junquera et al. 2020, Friis et al. 2016, Keovilignavong and Suhardiman 2020). There has also been some evidence of negative impacts due to increasing reliance on rubber production in the North, including for food security (Polthanee et al. 2021, Keovilignavong and Suhardiman 2020).

4.2.3.2 Rubber in Southern Laos

Production in the South is dominated by large-scale concessions, often involving international investors from Vietnam and Thailand (Smith et al. 2020). The Bolaven Plateau which extends across Attapeu, Champsack, Salavan, and Sekong is a major destination for foreign investment for rubber concessions (Hett et al. 2020). Southern rubber production has been associated with the disenfranchisement of local communities, particularly related to land and resource access and tenure (Özdoğan et al. 2018, Singh 2020). From a labor perspective, studies suggest that while these concessions do generate some local employment, more senior positions tend to be held by foreign workers (e.g. on Vietnamese-financed plantations), with a significant prevalence of illegal foreign labor (see Baird et al. 2019). There is also evidence that concessions cause negative environmental and social impacts, including deforestation, biodiversity loss, and food insecurity (Ahrends et al. 2015, Hurni and Fox 2018). Rubber concessions are one example of a broader trend of increased concessions in Laos since 2000, incentivized by the government policy framework of 'Turning Land Into Capital' (Ingalls et al. 2018, Hett et al. 2020, Friis and Nielsen 2016).

4.2.3.3 Rubber in Central Laos

Central production systems represent a blend of those used in the North and South, with some smallholder production and contract farming occurring, as well as a significant number of concessions (Douangsavanh, 2009, Smith et al. 2020). A study assessing land concessions in Vientiane and Bolikhamxai indicated negative socioeconomic impacts, including agricultural land loss for local communities and decreased access to non-timber forest products (NTFPs) (Yongnou et al. 2019). In terms of positive impacts, wage labor from plantations was reported to provide funds for education and medical costs (Ibid).

4.2.4 Rubber, Voluntary Sustainability Initiatives, and Economic Inequality

Due to its sustainability impacts, several efforts aim to mitigate the effects of rubber production, both in Southeast Asia and globally. This includes sustainable pilot plantation approaches, forest certification, agroforestry, government efforts, and roundtables like the Global Platform on Sustainable Natural Rubber (GPSNR) and Sustainable Natural Rubber Initiative (SNRI) (see Otten et al. 2020, Gitz et al. 2020, FSC 2021, PEFC 2021, Fair Rubber 2021). The GPSNR, a multi-stakeholder initiative including civil society groups and supply chain actors (e.g. traders, auto manufacturers, farmers) launched in 2018; this platform represents 50% of the global demand for natural rubber (GPSNR 2021). Since the GPSNR does not have a certification component, existing forest certification programs will likely be a key mechanism for companies to adhere to its requirements (see de Bonafos 2020).

These programs often acknowledge that the needs and challenges of smallholder and larger producers are different. Despite this, smallholder inclusion and participation are ongoing challenges for some voluntary programs, e.g. forest certification (Pinto and McDermott 2013, Higman and Nussbaum 2002, Karmann and Smith 2009). Additionally, these programs rarely explicitly address economic inequality in their activities, and consider broader equity concerns to varying degrees. For example, the SNRI is relatively sparse in terms of socioeconomic criteria, focusing primarily on child labor, forced labor, and freedom of association, but notably not addressing inequality or other social dynamics (IRSG 2014, Kenney-Lazar et al. 2018). Further, a previous review of four certification schemes – including Forest Stewardship Council (FSC) and the Programme for the Endorsement of Forest Certification (PEFC) - found that neither standard addressed overall distribution of costs and benefits from certification (McDermott 2013).

Despite these limitations, initiatives like the GPSNR have worked to incorporate inequality more explicitly. The GPSNR developed a policy framework to guide sustainable sourcing commitments from private sector members based on 12 principles; these principles emphasize aspects like tenure security, free prior informed consent (FPIC), grievance mechanisms, traditional resource access, and restitution for land appropriation, all of which are linked to inequality (GPSNR 2020). Additionally, some programs have placed a comparatively greater focus on procedural equity – for example via consultation with local communities – with distributive equity receiving less attention (Pinto and Mcdermott 2013, Pinto and McDermott 2012, Kalonga, Kulindwa, and Mshale 2015, McDermott 2013).

These varying degrees of focus on inequality pose an important limitation, because studies assessing previous crop booms and periods of monoculture expansion illustrate significant implications for income distribution (see Slothuus, Schmidt-Vogt, and Mertz 2020; Qaim et al. 2020, Cramb et al. 2017). Moreover, previous research suggests that rubber production can engender or exacerbate economic inequality (Evans et al. 2011, Phimmavong and Keenan 2020, Vongvisouk et al. 2014, Thongmanivong and Vongvisouk 2006). Based on previous research, we have identified several potential drivers of economic inequality important for rubber production as well as potential mechanisms for voluntary sustainability programs to engage and address these issues (see Table 4.1). These drivers are particularly relevant for the Laotian context but may also apply to other rubber production areas in the Greater Mekong and globally.

Table 4.1: Drivers of inequality for rubber production and potential mechanisms for engagement by voluntary sustainability initiatives.

Potential driver of inequality	**Potential mechanism to engage/address**	**Relevant citations (not exhaustive)**
Tenure insecurity	● Increase awareness of/engagement with participatory land use planning processes ● Incorporate baseline assessments and monitoring to ensure that program does not further entrench land inequalities ● Ensure environmental and social safeguards in place prior to new land investments/concessions ● Advocate for recognition of customary land claims	Kenney-Lazar 2012 and 2018; Ingalls et al. 2018; Suhardiman et al. 2015; Baird 2010; Junquera et al. 2020; Fox and Castella 2013
Dispossession from land and forest resources (including land grabbing and diminished NTFP access)	● Incorporate Free, Prior, Informed Consent (FPIC) requirements ● Include grievance mechanism ● Incorporate access to traditional forest resources ● Incorporate restitution for previous resource appropriation ● Ensure environmental and social safeguards in place prior to new land investments/concessions ● Advocate with policymakers to strengthen and enforce concessions regulations and contracts between foreign investors and local communities ● Support development of systems to address traceability and legality	Özdoğan et al. 2018; Singh 2020; Yongnou et al. 2019; Warren-Thomas, Dolman, and Edwards 2015; Smith et al. 2021; GPSNR 2020; Kenney-Lazar and Wong 2016; Baird 2010; Keovilignavong and Suhardiman 2020; Ingalls et al. 2018

Intra-household inequality (decision-making, labor) and potential disenfranchisement of female-headed agricultural households	● Incorporate gender equality baseline assessments and monitoring in the program ● Design specific mechanisms to engage female producers within the program (e.g. participation in training and employment)	Kusakabe and Chanthoumphone 2021; Kusakabe 2015; Loveridge et al. 2021
Insufficient worker protection/rights	● Ensure national and legal requirements for worker protections are met for all workers (e.g. ILO Conventions; minimum wage; collective bargaining) ● Ensure fairness of labor contracts and adherence to living wage ● Uphold occupational health and safety standards	Smith et al. 2020; Singh 2020; Kenney-Lazar 2012; Bennett 2018; Hett et al. 2020
Livelihood vulnerability due to rubber price/market fluctuations	● Facilitate livelihood diversification as a component of the program, for example via the production system itself (agroforestry) or through complementary approaches ● Advocate for price measures which buffer local communities, e.g. establishing or enforcing floor prices	Junquera and Gret-Regamey 2020; Cramb et al. 2017; Kenney-Lazar and Wong 2016; Vongvisouk and Dwyer, 2016; Jin et al. 2021
Barriers for smallholders to engage in supply chains and sustainability projects	● Design specific mechanisms to engage smallholder producers ● Ensure access to credit/technical expertise/extension ● Support formation of farmer associations and cooperatives for collective bargaining and sharing of best management practice/market information ● Support increasing value-added supply chain activities	Cramb et al. 2017; Kenney-Lazar and Wong 2016; Smith et al. 2020; Fox and Castella 2013; Vongvisouk and Dwyer, 2016; Pinto and McDermott 2013

4.3 Methods

4.3.1 Overview

This analysis combines novel land cover/land-use change (LCLUC) metrics based on remotely sensed imagery with nationally representative socioeconomic data to explore the relationship of rubber production to socioeconomic well-being in Laos, specifically through the lens of inequality decomposition and polarization.

4.3.2 Data sources and key variables

Key data sources include the widely used Laos Expenditure and Consumption Survey (LECS) for 2007/08 ('LECS 4') and 2012/13 ('LECS 5'), and village-level metrics describing rubber plantation development from 2000-2019, based on Landsat Analysis Ready Data (ARD, Potapov et al. 2020) and supervised machine learning classification.

4.3.2.1 LECS 4 (2007/08) and 5 (2012/13)

The LECS is a nationally representative household-level sample survey conducted over a period of twelve months and administered every five years since 1992/93 (Lao Department of Statistics 2010). The survey incorporates data regarding expenditures, consumption, production, investments, labor, nutrition, poverty, and income. The LECS 4 sample includes 518 villages, each with 15 to 16 households, with a total of 8,304 households (Lao Department of Statistics 2010). For LECS 5, the sample consists of 8,226 households and 515 villages. Half of the households surveyed for LECS 5 were also involved in the previous LECS sample, and the other half were randomly selected (Pimhidzai et al. 2014). The primary LECS variable of interest used in this analysis is our income measure, the monthly real consumption per capita of each LECS household, calculated by totaling household expenditures, minus the value of own-produced goods (e.g. food, firewood; National Statistical Centre 2004). Consumption is adjusted for annual price changes and reported in real values using 2003 as the base year. Notably, we exclude LECS households in predominantly urban areas from our analysis, as well as all villages in Vientiane Capital province, since we focus on rural economic inequality. It is important to note that because we only assess rural areas, we have eliminated arguably the largest driver of economic inequality and polarization within the country (see Epprecht et al. 2008); this enables us to discern trends in inequality among rural neighbors, rather than along rural and urban lines (which has been the focus of previous studies).

4.3.2.2 Village-level land-use areas

There are no nationally comprehensive official village boundaries available for Laos. To calculate rubber metrics that can be linked to specific villages, we utilize 2015 village-level land-use area polygons, a division of the total land area into individual village's likely land use space following the method of Epprecht et al. (2008), Messerli et al. (2008), and the 2015 Population and Housing Census. Polygons were calculated using GPS coordinates of the village centers from the 2015 Population and Housing Census using an equal travel time framework, considering factors like slope and road quality (Epprecht et al. 2008).

4.3.2.3 Rubber area estimation

We use a satellite-based rubber plantation extent data time-series to identify villages with significant areas of established plantations. It is challenging to map rubber plantations using medium resolution (30 m/pixel) remotely sensed data, particularly vis-a-vis distinguishing rubber from secondary forests or other tree crops (see Hurni and Fox 2018). Here, we employed a supervised classification to map rubber plantations as a land cover class following the approach of Potapov et al. (2019) and use the Landsat ARD 2000-2019 time-series data as classification input (Potapov et al. 2020). The calibration data (rubber plantation polygons) were manually collected through visual interpretation of satellite images and informed by land cover maps from Hurni and Fox (2018) and Özdoğan, Baird, and Dwyer (2018). The classification objective was to indicate the presence of mature rubber plantations. Young (< 5 years) plantations usually have no distinct spectral characteristics that can be used to consistently separate them from other plantation types (Dong et al. 2013).

We calibrated and applied the rubber plantation mapping model for the year 2019. The rubber map validation using a stratified random sample of 500 Landsat pixels showed an overall map accuracy of 99.6% (SE 0.2%), producer's accuracy (which indicates the rubber omission) of 68.8% (SE 10.8%), and user's accuracy (which indicates the rubber commission) of 94% (SE 3.4%). The sample-based rubber plantation area estimate is 27% larger than the mapped area which confirms that the map omits a portion of rubber extent. Visual analysis suggests that our mapping algorithm has a higher omission rate for recently established and small-scale rubber plantations, particularly if they are intermixed with other woody crops. However, the map correctly reflects the location of most rubber plantations and is suitable for our national-scale analysis. Future local-scale analyses may require higher precision of rubber plantation mapping using higher spatial and temporal resolution satellite data. We applied the same model to the years 2003, 2005, 2010 and 2015. The annual rubber plantation maps were integrated into a single map reflecting the date of plantation establishment. Following the time frame of the LECS socioeconomic data, only villages with more than 1 ha of rubber area detected between 2003 and 2010 were considered rubber production areas for the purposes of the analysis.

4.3.2.4 Village matching

In both the LECS and the village-level polygon data, villages are identified using unique numeric IDs. Between the time of LECS 4 (2007/08) and the 2015 Population and Housing Census, several administrative processes resulted in changes to village composition in Laos. This included village resettlement, merging and divisions of existing villages, and changes in ID codes (Epprecht et al. 2018). Therefore, to conduct longitudinal analysis, village land-use areas must be matched with their counterparts in other years, despite the possibility of dissimilarities in numeric IDs (as well as the population residing in the village area). Here, we achieved this by first matching LECS 4 and 5 villages to 2015 land-use area polygons using the official census village ID. For households not successfully matched using village ID (approximately 20% of the households), we conducted a one-to-one spatial match in QGIS to associate their LECS village coordinates to the 2015 village land-use area polygons.

4.3.3 Analytical approach

4.3.3.1 Assessing inequality through the Gini coefficient, Thiel L and Thiel T

We assess economic inequality primarily through calculation and decomposition of the Gini coefficient. We also present results based on the Thiel L and T indices. For all three indices, higher values indicate greater inequality.

4.3.3.2 Decomposing the Gini coefficient

Gini decomposition enables an improved understanding of comparative within and between-group contributions to inequality, and is an important tool in evaluating drivers and changes in socioeconomic well-being and informing policy design (Silva, Matyas, and Cunguara 2015, Araar 2006). The Gini coefficient is arguably the most popular measure of income inequality, with a value ranging from 0 (total equality) to 1 (total inequality) (Chen, Tsaur, and Rhai 1982, Correa-Parra, Vergara-Perucich, and Aguirre-Nuñez 2020). The Gini decomposition is illustrated with the following formula:

$$I = \sum_{g=1}^{G} \emptyset_g \varphi_g I_g + \bar{I} + \bar{R}$$

Where I = the total Gini index. The first term reflects within-group inequality, with G = the total population of the subgroups, $\emptyset_g$ = population share of group g, φ_g = income share of group g, and I_g = Gini index for group g. $\bar{I}$ = Between-group inequality, and $\bar{R}$ = the residual from overlapping income levels across groups, also known as overlap inequality.

The Gini decomposition was calculated using the "gini decomp" function in R package *dineq* (Schulenberg 2018, Mookherjee and Shorrocks 1982, Cowell 2000). We compare Gini decompositions by rubber group and by geographic region. We classify villages according to the presence of rubber (yes or no, >1 hectare) and geographic region for a total of six groups. Additional analyses decomposing inequality by ethnic group can be found in the Appendix (Section A3.2).

4.3.3.3 Duclos Esteban Ray (DER) Index

The DER polarization index measures the extent to which similar incomes cluster along an income distribution and can be decomposed in a similar manner to the Gini index (Duclos, Esteban, and Ray 2004). The DER coefficient values also range from 0 to 1, and higher values denote a greater degree of polarization.

The DER index can be decomposed as follows:

$$P(\alpha) = \underbrace{\sum_{g=1}^{G} \phi_g^{1+\alpha} \varphi_g^{1-\alpha} R_g P_g}_{Within} + \underbrace{\overline{P}}_{Between}$$

where

$$R_g = \frac{\int a_g(x)\pi_g(x)f(x)^{1+\alpha}dx}{\varphi_g \int a_g(x)f_g(x)^{1+\alpha}dx}$$

G = Total population of subgroups g; α = A normative 'inequality aversion' parameter that expresses the sensitivity of the index to inequality ϵ [.25-1]; ϕg = Population share of group g; Φg = Income share of group g; Pg = DER index for group g; πg = Local proportion of households belonging to group g and having income x; f(x) 1+α = Density function of income x; fg(x) 1+α = Density function of income x for group g; ag(x) = The absolute distance from income x and other incomes in group g; dx = Differential distance or spread of income x from the median.

Following the method outlined in Silva et al. (2015) and Araar (2008), we also use the components of the DER index to estimate the magnitude of relative deprivation and surplus income among households within each decomposition group. If a particular group is composed of a large proportion of poor households relative to the group mean, the ratio Dg/Sg (i.e. income deficit to income surplus) will be relatively higher compared to groups with larger concentrations of higher income households. This ratio allows us to analyze changes in polarization based on the direction of income shifts over time along the distribution.

The DER assesses trends in the income distribution and can illustrate the extent to which polarization changes over time. However, it is not best-suited to explain how exogenously defined groups contribute to this polarization (see Indra et al. 2018). Thus, we also conduct a Zhang-Kanbur (Z-K) index analysis, which can be found in the Appendix (A3.4).

4.3.3.4 Limitations

Changes in rubber production systems were not the only driver of rural economic inequality over the study period. For example, land concessions for agriculture, other tree plantations, and mining may have altered livelihoods composition and resource access (Hett et al. 2020). We do not control for pre-existing differences in the rubber and nonrubber groups, or other factors over the study period that could have significantly impacted economic inequality. We address this by exploring comparative changes among other groups, e.g. regional dynamics and ethnic composition, and incorporating a discussion of potential alternative explanations for our results (see Section 4.5.1.4). Further research is needed to investigate these dynamics, for example via field surveys and/or through statistical approaches to control for differences between rubber and non-rubber producing areas.

The study timeframe also presents some important considerations and limitations. Previous research suggests that some Northern rubber trees were likely more mature as compared to Southern and Central rubber plantations during the study period (Epprecht et al. 2018, Evans et al. 2011, Özdoğan, Baird, and Dwyer 2018). For this reason, Northern rubber production areas may have seen greater increases in household income from 2007/08-2012/13 due to rubber sales, whereas in the South, there may have been a greater reliance on pre-harvest activities, for example clearing, planting, and weeding via wage labor.

Finally, there are important aspects of inequality that fall outside the scope of this analysis. This includes intra-household dynamics and gender inequality (e.g. whether inequality unfolds differently for female-headed households, or the extent to which rubber adoption affects labor divisions and decision-making within households). These topics present possibilities for future research.

4.4 Results

4.4.1 Descriptive statistics

Before turning to the results of our economic inequality and polarization analyses, we provide descriptive statistics on rubber production and household consumption by region. We divide households into six groupings: Northern Laos with rubber (Group 1), Northern Laos without rubber (Group 2), Central Laos with rubber (Group 3), Central Laos without rubber (Group 4), Southern Laos with rubber (Group 5), and Southern Laos without rubber (Group 6). Additional information on characteristics of our study sample groupings (e.g. ethnic composition, elevation, and land access) is provided in the Appendix (Tables A3.3-A3.6). Overall, the rubber groups represent a smaller set of provinces than their non-rubber counterparts. This includes Louangnamtha and Bokeo (Group 1); Vientiane, Bolikhamxai, Khammouan, Savannakhet, and Xaisomboon (Group 3); and Salavan, Champasack, and Attapeu (Group 5).

4.4.1.1 Rubber production area

Based on our rubber plantation area estimates, we identify production area increases in the rubber-growing[20] LECS villages from 2003 to 2015 (Table 4.2). We see the largest increase in new rubber detections from 2010 to 2015 – which, given the lag in detection time, comports

[20] As shorthand, we refer to groups as "rubber producing" and "non rubber producing" or "with rubber" and "without rubber", but it should be noted that these are more accurately considered groups in areas *with rubber being produced in the village as a whole* (during the study period) and groups *without rubber being produced in the village,* since not every household in the village produces rubber.

with previous research suggesting that the bulk of the rubber boom occurred between 2006 and 2011 (Hurni and Fox 2018). Overall from 2003-2010, Group 1 exhibited 464 hectares of new detections, Group 3 showed 574 hectares of new detections, and Group 5 showed 542 hectares of new detections. A map of rubber detections from 2003-2010 is displayed in Figure 4.2.

Table 4.2: Rubber production area

Description: Estimated baseline and newly detected rubber plantation area by rubber production groups in the study area. The first column shows the percentage of village area, and the second column shows the area in hectares. The mean values reflect all villages in the group. *Data*: Rubber plantation area estimates based on Landsat ARD and supervised classification.							
		Group 1: Northern Laos, with rubber *Village n = 5* *Household n = 80*		**Group 3: Central Laos, with rubber** *Village n = 9* *Household n = 144*		**Group 5: Southern Laos, with rubber** *Village n = 4* *Household n = 64*	
		% Total Area	**Hectares**	**% Total Area**	**Hectares**	**% Total Area**	**Hectares**
Rubber area in 2003 (*base year*)	mean	0.04%	1.44	0.01%	0.28	0.03%	2.22
Newly detected rubber area* 2003-2010	mean	0.65%	5.80	0.18%	4.00	0.21%	8.48
Newly detected rubber area 2010-2015	mean	5.79%	75.07	1.55%	81.45	2.01%	50.21
Total estimated rubber area 2003-2015	mean	6.48%	82.31	1.65%	85.71	2.24%	60.91
Average village land area (hectares)	mean		2,837		4,913		4,874

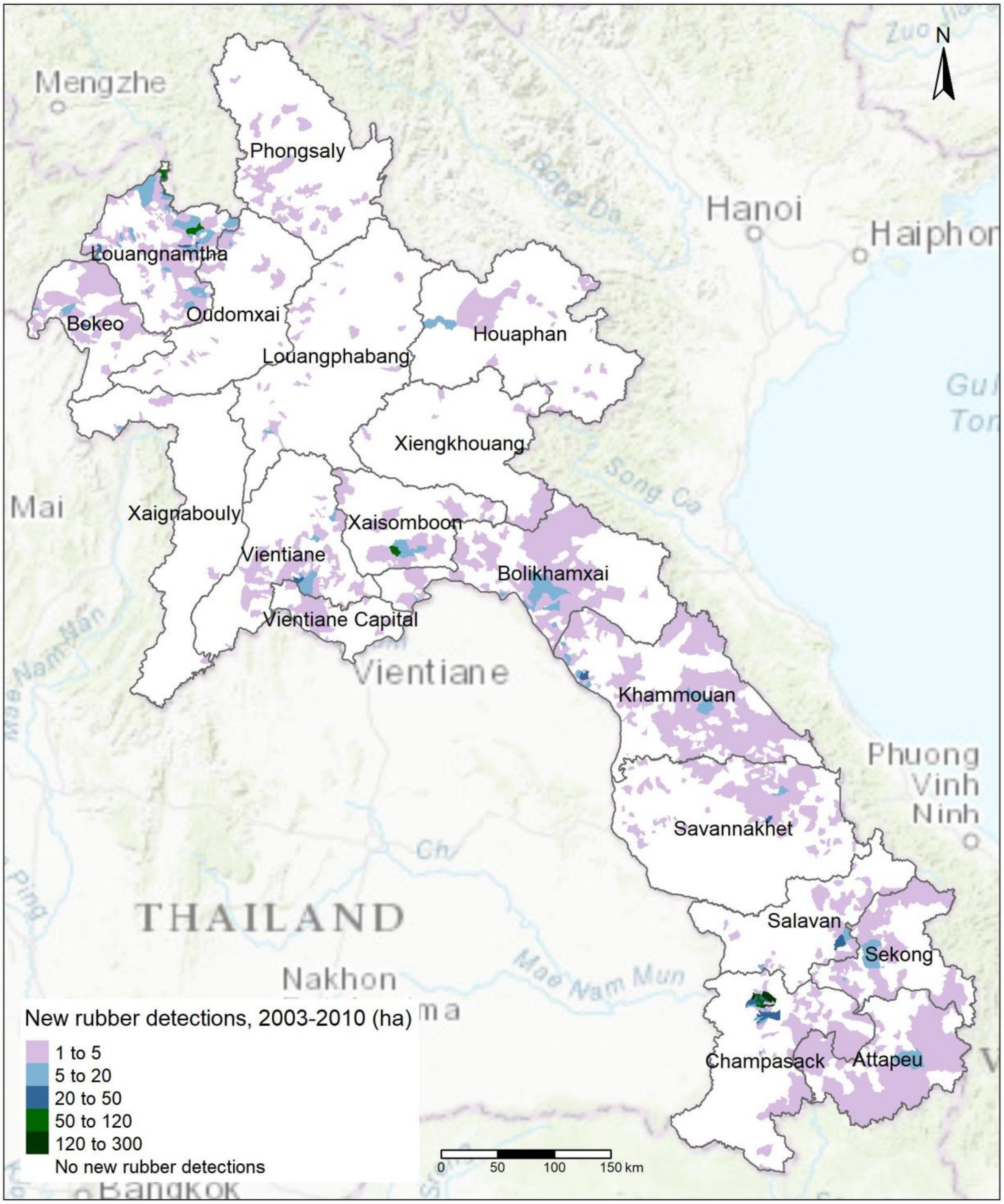


Figure 4.2: Village-level map of rubber area. From 2003-2010, new rubber detections were highest in Champasack (39% of new detections, approximately 1,237 hectares) and Louangnamtha (27%, 932 hectares), followed by Khammouan (8%) and Bolikhamxai (4%). Data: Authors' elaboration based on (1) data from Open Development Mekong (2021) and Esri Topographic map; (2) Laos village polygons (Epprecht et al. (2008), Messerli et al. (2008), and the 2015 Laos Population and Housing Census); and (3) rubber plantation area estimates based on Landsat ARD and supervised classification.

4.4.1.2 Consumption per capita

Overall, mean real consumption per capita increased by 133% from 2007/08 (152,850 kip) to 2012/13 (356,675 kip) and all three regions experienced substantial gains. This reflects broader improvements in poverty reduction over the study period (Pimhidzai et al. 2014). See Appendix (Section A3.1) for additional detail. As shown in Figure 4.3, initial mean per capita consumption in 2007/08 was lowest in the Central region (approximately 147,450 kip), followed by the South (154,470 kip), and then the North (157,600 kip). In 2012/13, the South had the highest observed mean per capita consumption (377,860 kip), followed by the Central region (359,430 kip), then the North (341,320 kip). These changes represent growth rates of 117% in the North, 144% in the Central region, and 145% in the South.

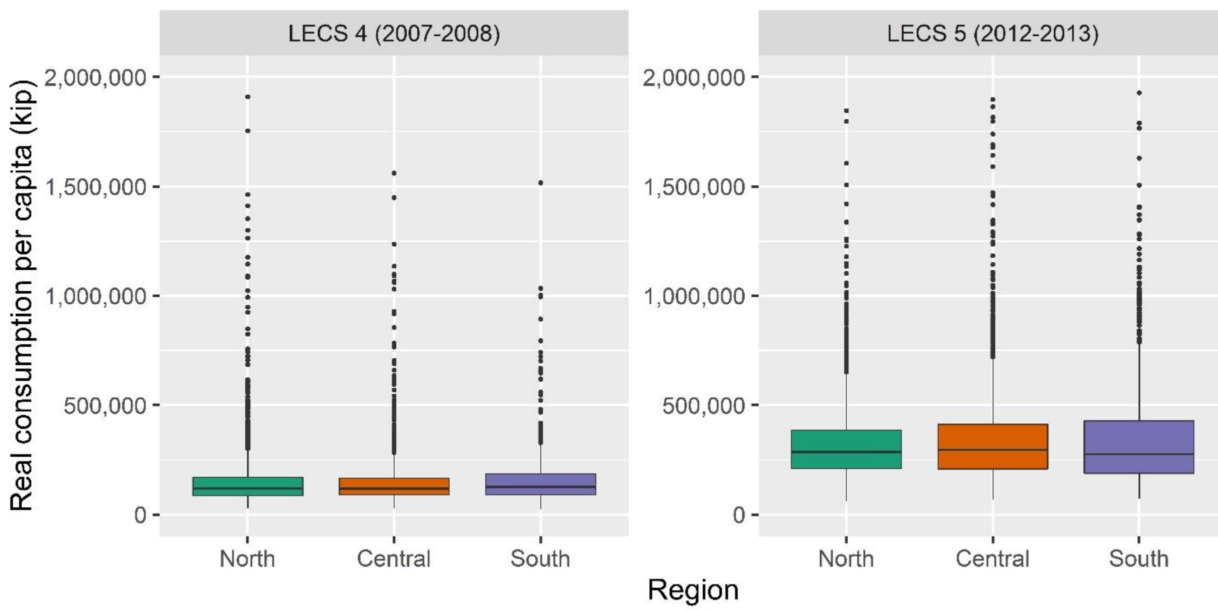


Figure 4.3: Monthly real consumption per capita by region (rural areas), 2007/08-2012/2013. Source: Authors' estimates using the LECS4 and LECS5 Household Survey Data adjusted to 2003 values. Outliers over 2,000,000 kip removed for visualization purposes.

When we explore income by regional rubber groupings, we observe the greatest gains in mean per capita consumption in the Northern and Central rubber-producing groups, with increases of 192% and 189% respectively (Groups 1 and 3, see Table 4.3). Figure 4.4 illustrates these trends visually through a boxplot. This demonstrates that the Northern and Central with-rubber groups exhibited mean per capita consumption increases well over the average for their respective regions. In contrast, consumption figures in our two Southern groups (i.e. with and without rubber) experienced comparable gains, with growth in the rubber production group slightly lower than its regional counterpart. Additional information on real consumption per capita for various ethnic groups can be found in the Appendix (Section A3.2).

Table 4.3: Monthly real consumption per capita by rubber group, 2007/08-2012/13 (kip).

	LECS 4 (2007/08)		**LECS 5 (2012/13)**		
	mean	**median**	**mean**	**median**	**% change in mean**
Group 1: Northern Laos, with rubber	143,380	121,646	418,540	368,290	191.9%
Group 2: Northern Laos, without rubber	158,390	119,114	337,290	282,390	113.0%
Group 3: Central Laos, with rubber	148,780	123,616	429,380	341,310	188.6%
Group 4: Central Laos, without rubber	147,340	118,824	354,140	293,480	140.4%
Group 5: Southern Laos, with rubber	132,220	119,435	321,200	208,060	142.9%
Group 6: Southern Laos, without rubber	155,860	126,084	381,150	280,920	144.6%
Source: Authors' estimates using the LECS4 and LECS5 Household Survey Data.					

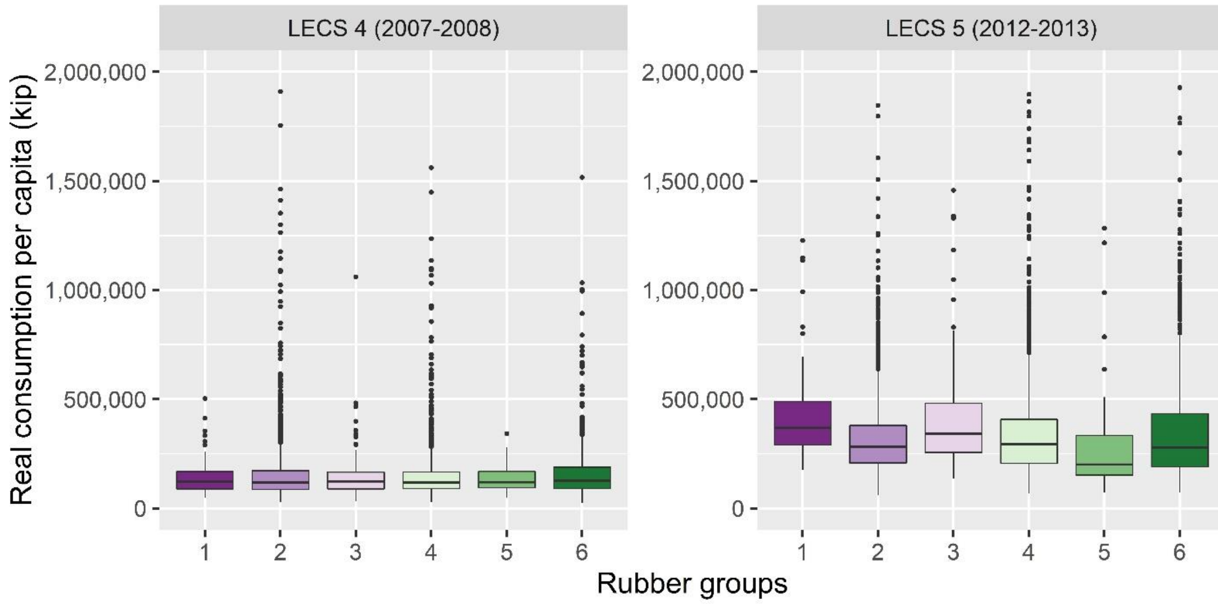


Figure 4.4: Monthly real consumption per capita by rubber group, 2007/08-2012/13 (kip). Source: Authors' estimates using the LECS4 and LECS5 Household Survey Data. Outliers over 2,000,000 kip removed for visualization purposes.

4.4.2 Inequality Metrics and Gini Decomposition

We present trends in rural economic inequality from 2007/8-2012/13 using the Theil L, Theil T, and Gini coefficient measures (Table 4.4). All three measures exhibit very slight reductions, with the Theil T measure exhibiting the largest absolute decrease. Previous studies using LECS data have shown relatively stable rural inequality in Laos from 2007/08 to 2012/13, although longer-term trends since the 1990s indicate increasing inequality (Warr, Rasphone, and Menon 2015, Pimhidzai et al. 2014).

Table 4.4: Overall rural inequality trends, 2007/08-2012/13.

	LECS 4 (2007/08)	LECS 5 (2012/13)	Absolute change
Theil (L) / GE (0)	0.185	0.178	-0.007
Theil (T) / GE (1)	0.234	0.217	-0.017
Gini coefficient	0.336	0.333	-0.003
Source: Authors' estimates using the LECS4 and LECS5 Household Survey Data.			

We first decompose the Gini coefficient by region only (Table 4.5), and then based on rubber production status and geographic region (Table 4.6). When decomposing the Gini coefficient by region, we find inequality rises in the South and Central areas of the country, by 17% and 2% respectively, and decreases in the North by 15% (Table 4.4). Our finding for rural economic inequality diverges slightly from previous research which looked at the longer-term trend in regional inequality from 1992 to 2013 in both rural and urban areas, finding an overall increase in inequality throughout Laos (Warr, Rasphone, and Menon 2015). This is likely caused by our omission of urban households from this analysis.

When we assess economic inequality by rubber production status and region, we find that inequality decreased overall for both rubber-producing and non-rubber producing village groups in the North. Inequality decreased to a slightly larger extent in the 'without rubber' groups by about 16%, as measured by the overall Gini coefficient. Similarly, in the Central region, inequality increased for both rubber-producing and non-rubber producing groups, with a greater increase in inequality in rubber-producing villages of about 7%. In the Southern region, we see a comparatively large increase in inequality for the rubber-producing group, of 66%. We also observe an increase in inequality in the non-rubber producing groups in the South of about 16%. Overall, within-group inequality remained constant over the 5-year study period, suggesting that inequality among neighbors did not change. Between-group inequality decreases slightly, indicating that incomes become more similar across all groups. Overlap inequality is the largest contributor to the overall Gini coefficient in both years, indicating a significant degree of similarity in income ranges among the groups; within-group inequality is the second largest driver of inequality.

Table 4.5: Gini index decomposition for rural per capita consumption by region, LECS 4 (2007/08) and LECS 5 (2012/13).

	Population share	**Gini coefficient**	**Population share**	**Gini coefficient**	**% change**
Region	**LECS 4 (2007/08)**		**LECS 5 (2012/13)**		
North	0.342	0.364	0.321	0.306	-15.9%
Central	0.419	0.311	0.43	0.318	2.3%
South	0.239	0.327	0.249	0.383	17.1%
Between-group inequality		0.031		0.022	-0.9%
Within-group inequality		0.115		0.114	-0.1%
Overlap inequality, full sample	0	0.19	0	0.197	3.7%
Gini coefficient, full sample	100%	0.336	100%	0.333	-0.9%
Source: Authors' estimates using the LECS4 and LECS5 Household Survey Data.					

Table 4.6: Gini index decomposition for per capita consumption by rubber production groups, LECS 4 (2007/08) and LECS 5 (2012/13).

Rubber group & geographic region	Population share	Gini coefficient	Population share	Gini coefficient	% change
	LECS 4 (2007/08)		**LECS 5 (2012/13)**		
Group 1: Northern Laos, with rubber	1.3%	0.273	1.3%	0.244	-10.6%
Group 2: Northern Laos, without rubber	32.8%	0.367	31.0%	0.307	-16.4%
Group 3: Central Laos, with rubber	4.0%	0.297	2.7%	0.319	7.4%
Group 4: Central Laos, without rubber	38.0%	0.312	40.0%	0.317	1.6%
Group 5: Southern Laos, with rubber	1.2%	0.221	1.0%	0.367	66.1%
Group 6: Southern Laos, without rubber	22.7%	0.329	24.0%	0.382	16.1%
Within-group inequality, full sample		0.102		0.102	0.0%
Between-group inequality, full sample		0.037		0.033	-10.8%
Overlap inequality, full sample		0.197		0.199	1.0%
Gini coefficient, full sample	100%	0.336	100%	0.333	-0.9%

Source: Authors' estimates using the LECS4 and LECS5 Household Survey Data.

4.4.3 DER Index

Similarly to our results for economic inequality, we find regional variation in polarization according to rubber production status as measured by the DER (Tables 4.7 and 4.8). As shown in Table 4.8, polarization declines in Groups 1 and 2 and increases in the remainder of the groups. We also find that rubber-production areas make less progress in reducing polarization than their regional non-rubber counterparts; polarization increases more in the southern with-rubber group (Group 5) and declines less in the Northern (Group 1) and Central (Group 3) regions. The initial DER index values from 2007/08 are in a similar range across the groups, with the 'no rubber' group in the North (Group 2) having the highest initial polarization index. However, more variation across groups exists by 2012/2013. Changes in the D/S ratio suggest that Groups 1, 3, and 4 gained more households on the higher end of the income distribution, while Groups 5 and 6 gained more households on the lower end of the distribution.

Table 4.7: DER decomposition of total household consumption per capita by region (alpha=.5).

Geographic Region	**DER Coefficient**	**D/S Ratio**		**DER Coefficient**	**D/S Ratio**
	2007/2008			**2012/2013**	
North	0.238	2.356		0.208	2.254
Central	0.210	2.569		0.213	2.158
South	0.218	1.776		0.244	2.070
Within-Group Polarization for Full Sample	0.076			0.074	
Between-Group Polarization for Full Sample	0.144			0.144	
DER Coefficient for Full Sample	0.220	2.280		0.218	2.166
Source: Authors' estimates using the LECS4 and LECS5 Household Survey Data.					

Table 4.8: DER decomposition of total household consumption per capita by region and rubber production status (alpha=.5).

Rubber Group & Geographic Region	**DER Coefficient**	**D/S Ratio**		**DER Coefficient**	**D/S Ratio**
	2007/2008			**2012/2013**	
Group 1: Northern Laos, with rubber	0.221	2.503		0.202	1.113
Group 2: Northern Laos, without rubber	0.242	2.352		0.209	2.323
Group 3: Central Laos, with rubber	0.209	2.863		0.233	1.438
Group 4: Central Laos, without rubber	0.212	2.604		0.213	2.223
Group 5: Southern Laos, with rubber	0.218	2.794		0.266	4.467
Group 6: Southern Laos, without rubber	0.218	1.762		0.244	2.002
Within-Group Polarization for Full Sample	0.061			0.066	
Between-Group Polarization for Full Sample	0.159			0.152	
DER Coefficient for Full Sample	0.220	2.280		0.218	2.166
Source: Authors' estimates using the LECS4 and LECS5 Household Survey Data.					

4.5 Discussion

4.5.1 Economic inequality and polarization trends

At the national level, the analysis results support previous findings of relatively stable rural economic inequality in Laos between 2007/8 and 2012/13. However, this trend masks substantial regional changes. In Northern Laos, inequality decreases over the study period, but this progress is offset at the national level by widening income gaps in the Central and Southern regions. Our findings for overall rural polarization exhibit a similar pattern of stability at the national level, with regional variation. Northern Laos again demonstrates different distributional dynamics with polarization (as defined by the DER index) trending downwards while the opposite occurs in the Center and South. Indeed, polarization in the South increases by almost as much as it declines in the North. Although the mean income per capita increased in every region, according to the D/S ratio, Northern and Central Laos had a larger proportion of poor households relative to the group mean in 2007/08 than in 2012/13. The opposite holds true for the South, which gained a larger concentration of lower income households by the end of the study period.

Our findings indicate that distributional dynamics differ between areas with and without rubber, and that rubber production is associated with increased economic inequality and polarization. This comports with previous research conducted elsewhere in the Greater Mekong region (see Jin et al. 2021). Baseline inequality levels were lower in 2007/08 for rubber production groups in every region (as compared to their non-producing regional counterparts). By 2012/13, inequality is still lower in Northern and Southern rubber production groups but the gap between with- and without-rubber areas shrinks. In the Central region, inequality becomes roughly equal between the two groups. Groupings in the same region always trend in the same direction but the magnitude of change – as measured by the Gini coefficient and DER index – differs by rubber production status.

Variations in rubber production systems are likely a key driver in these regional differences. We find the most extreme differences occurring between the North and South, where rubber production systems take distinctly different forms (largely smallholder driven in the North, and concession-driven in the South). In the North, where the distribution improves, inequality and polarization decrease less in rubber production areas (Group 1). In the South, where the distribution worsens, we find increases in inequality and polarization are much larger in the with-rubber grouping (Group 5). For example, Group 1 shifts from one of the highest polarization values to the lowest based on the DER results, whereas Group 5 experiences the greatest increase and becomes the most polarized by 2012/13. Notably, our Z-K index results (Appendix A3.4) suggest that there may be additional factors driving polarization among the groups aside from rubber production.

The revenues in smallholder and concessions systems are distributed much differently, with companies retaining the revenues in the concession model, and local people involved primarily via wage labor. Another potential driver of these differences is variations in rubber plantation age during the study period, with a higher prevalence of more mature plantations in the North (Epprecht et al. 2018). This likely contributed to higher rubber productivity in the North during the LECS 5 data collection period (2012/13) as compared to the South. Rubber in earlier productive stages requires less labor for clearing, planting, and weeding as compared to harvesting, and brings lower wages and revenues (see Evans et al. 2011, Baird 2010).

Previous studies have found that farmers with greater initial access to capital and land may be more likely to benefit from rubber transitions, whereas poor households are more likely to become wage laborers for cash crop farmers and see fewer benefits (Vongvisouk et al. 2014, Thongmanivong and Vongvisouk 2006; Cramb et al. 2017). In the Center and South, research has suggested that concessions may alter or diminish communities' access to land and resources, and foster labor dynamics with negative distributional outcomes (Yongnou et al. 2019, Özdoğan et al. 2018, Singh 2020). Below, we discuss regional trends separately and highlight the implications of our findings for voluntary sustainability programs, both in Laos and more broadly.

4.5.1.1 Northern Laos

Our results suggest that Northern Laos experienced the most progress from an economic inequality perspective, as rural consumption became more equal and less polarized for both rubber and nonrubber producing groups. The highest gains in per capita consumption were found in Group 1. This group also experienced decreased inequality and polarization, although less so than Group 2, its regional non-rubber counterpart. The declining D/S ratio suggests that in Group 1, rising incomes were more evenly spread along the distribution than in Group 2.

There are a few ways in which smallholder rubber systems may have contributed to this progress. First, previous research has suggested that rubber production expanded in the North through social and investor networks transmitting price information, inputs, and other resources, which may have enabled more farmers to participate in smallholder rubber production than in concession-based models (see Junquera et al. 2020, Baird and Vue 2017). Revenues in smallholder production systems also stay with the producers, unlike in concessions. In some cases, producers organized in Rubber Grower Associations or other networks to negotiate prices and contracts (Douangsavanh, Thammavong, and Noble 2008, Fox and Castella 2013). Group 1's proximity to the Chinese border and high proportion of Chinese-Tibetan and Mon-Mien households likely enabled greater uptake of rubber production and associated economic benefits (see Appendix A3.3, Phimmavong and Keenan 2020).

In some cases, smallholder production allowed households to diversify and grow rubber on their own plantations but continue farming other crops. Studies have shown that this diversification resulted in better well-being outcomes as compared to complete reliance on rubber (Polthanee et al. 2021, Kenney-Lazar and Wong 2016). Further, rubber investments for smallholders brought co-benefits in some cases, e.g. increased land tenure security (Smith et al. 2020). This may be reflected in our LECS sample, since Group 1's mean area of land ownership increases over the study period, while Group 2's mean area of land decreases (Appendix Table A3.6). Finally, contract farming arrangements – which are particularly common in Phongsaly, Bokeo, and Oudomxai – involve benefit sharing with investor companies (Smith et al. 2020). This may have decreased economic gains as compared to non-contract farming rubber households, and thus moderated increases in local inequality. However, the greater reduction in economic inequality in the nonrubber villages in the North suggests that rubber is still associated to some extent with 'winners' and 'losers'; i.e., the better-off farmers made larger initial investments in rubber, thus reaping a greater portion of the benefits (see Evans et al. 2011, Hepp, Bech Bruun, and de Neergaard 2019, Jin et al. 2021).

It is important to note that previous studies have indicated some negative livelihood impacts of rubber production in the North, including in cases of land concessions by Chinese companies (e.g. Friis et al. 2016, Keovilignavong and Suhardiman 2020). These impacts may not be reflected in our results due to the predominance of individual and contract farming in Louangnamtha and Bokeo, where our Group 1 households are located. Additionally, economic inequality and poverty trajectories may have shifted in recent years due to drops in the rubber price; i.e, the gap between rubber and non-rubber producing households may have narrowed (see Jin et al. 2021). Previous research finds that rubber producing households adapted to decreasing prices through a variety of measures, including waiting to tap trees, relying upon only household labor, bargaining collectively for slightly higher prices, selling or leasing plantations, planting other crops, and pursuing off-farm livelihoods (Vongvisouk and Dwyer 2016, Jin et al. 2021). Several studies have noted that rubber producing households with less diversified livelihoods and a higher dependence on rubber are more vulnerable to price changes, underlining the need for government price support, household income diversification strategies, and development of the rubber processing industry, among other measures (Polthanee et al. 2021, Jin et al. 2021). The impact of price changes on rubber-producing households' livelihoods is an area for continued research.

4.5.1.2 Southern Laos

Our results indicate a troubling trend of increasing rural economic inequality and polarization, in the South generally, and for rubber-producing villages in particular. These findings match previous research suggesting increased inequality in the South over this period (Ingalls et al. 2018, Warr, Rasphone, and Menon 2015). Although inequality experiences a slight increase in Group 6 (without rubber), the substantial increase in inequality in Group 5 suggests that differential vulnerability is greater among households in rubber production areas. It is likely that these dynamics are driven – at least in part – by negative impacts of rubber concessions, including dispossession from land and forest resources (Keovilignavong and Suhardiman 2020, Özdoğan et al. 2018, Singh 2020, Kenney-Lazar 2012, Baird 2010). Group 5 is comprised of households from provinces dominated by concessions – in Salavan and Attapeu concessions represent virtually all rubber production, whereas Champsack is primarily concession-based with a small amount of contract and individual farming (Smith et al. 2020).

Labor dynamics, which differ significantly between smallholder production and concessions systems, may also influence our results. In concessions, studies report that laborers have found tapping income to be insufficient, particularly after price decreases in 2011 (Baird and Fox 2015). Lack of contracts, piece-rates, and seasonal pay can also cause negative economic outcomes for tappers - and there have been concerns that concession and contract models do not provide sufficient employment opportunities for local people (Baird 2010, Smith et al. 2020). Furthermore, a recent study found that better paid jobs in concessions are typically not filled by locals (see Nanhthavong et al. 2022, *forthcoming*).

We also observe increased polarization as measured by the DER index in the South, accompanied by a greater number of households on the lower end of the income distribution, suggesting the potential for social tension and conflict. The role of rubber concessions as a driver of social conflict is well-established in the literature, and our analysis supports this relationship (Nanhthavong et al. 2020, Baird 2010; Baird 2020). Some studies suggest a link between land dispossession and increased poverty in rubber production areas, for example in the Bolaven plateau in Champasack and Salavan (Nanhthavong et al. 2020, Ingalls et al. 2018). The poorest households have also been found to suffer most under land dispossession (Nanhthavong et al. 2022, *forthcoming*). This dynamic may be reflected in our rubber villages in the South – during the study period, Group 5 experienced a decrease in the mean land area owned from 2.32 ha to 1.62 ha, the largest decrease of all groups (Appendix, Table A3.6). Moreover, research suggests that in many concession-affected villages, only a small number of households (mostly poorer households) work on concessions, while higher-income households have better income alternatives (Nanhthavong et al. 2022, *forthcoming*). Other research suggests that because the Land and Forest Allocation program was implemented earlier in the North, Northern rubber villages may have benefited from comparatively secure tenure relative to Southern villages (Cramb et al. 2017).

Finally, the rubber and nonrubber groups in the South also have different ethnic compositions, which could contribute to differences in economic inequality and polarization outcomes. Notably, Group 6 has a larger portion of households in the Lao-Tai group, while Group 5 has a larger portion of households in the Mon-Khmer group (Appendix Table A3.4). The Mon-Khmer face higher levels of poverty as compared to the Lao-Tai, as well as larger increases in inequality from 2002-2013 (Bader et al. 2017, Warr, Rasphone, and Menon 2015).

4.5.1.3 Central Laos

In Central Laos, we observe a few key trends: first, per capita consumption increases over the study period, with higher increases in the with-rubber group (Group 3). Inequality increases across rural areas in the Central region generally, with a larger increase in the with-rubber group. However, the increases are not as large as the trends we observe in the South. Polarization as measured by the DER increases slightly in the Central region overall, and the D/S ratio suggests that more households moved to the higher end of the income distribution. Larger improvements in the D/S ratio are observed in the with-rubber Group 3. Taken together, these trends indicate that rubber production is associated with increases in per capita consumption, and slight increases in inequality and polarization in the Central region. So, we see some of the positive outcomes of the North from an income perspective, and some of the negative outcomes of the South from an inequality and polarization perspective.

There are five Central provinces represented in Group 3. These areas represent a variety of production systems – Vientiane and Bolikhamxai are home to a mix of individual farmers and concessions, Khammouan is dominated by individual farmers, and Savannakhet is primarily concession-based (Smith et al. 2020). Previous research has suggested that concessions in Central Laos bring some of the negative impacts observed in other regions, including loss in access to land and resources (Yongnou et al. 2019, Smith et al. 2020). However, there is also evidence that wage labor has brought income benefits, and a 2020 report suggested that overall, Central concessions performed better on environmental and social criteria than Southern concessions (Yongnou et al. 2019, Hett et al. 2020).

4.5.1.4 Potential alternative explanations and broader trends

Other major economic drivers and significant land use changes may have, in addition to rubber production, shaped changes in rural consumption and economic inequality over the study period. One key example is other large-scale land concessions (Hett et al. 2020, Ingalls et al. 2018). During the study period, the largest areas for land concessions were granted for mining (zinc/tin, copper), rubber, eucalyptus, mixed plantations (e.g. rubber/teak, rubber/acacia), sugarcane, coffee, and jatropha (Schönweger et al. 2012). Notably, though, rubber has consistently been one of the most common crops for land leases and concessions; a 2012 study found that it was the most prominent tree crop planted via concession or lease, with approximately 130,000 ha representing 225 separate investment projects (Schönweger et al. 2012).

Overall, the rise of agricultural commercialization in Laos has been accompanied by concerns related to rural indebtedness, landlessness, and food security (see Schönweger et al. 2012, Ingalls et al. 2018). Additionally, previous research has posited that an increase in mining investments in the South in recent years has increased inequality (Warr, Rasphone and Menon 2015). However, it is notable that these transitions are occurring in rubber producing areas. Therefore, it is important for sustainability initiatives – both those that are already implementing activities, and those in development – to consider the importance of these changes for sustainability outcomes.

4.5.2 Implications for voluntary sustainability programs

Sustainability programs for rubber have cropped up to address the clear and multifaceted linkages between rubber production, socioeconomic well-being, and environmental health. In Laos in particular, rubber is often produced by communities facing poverty and marginalization and grown in previously forested land, underscoring the importance of sustainability interventions (Kenney-Lazar et al. 2018, Schönweger et al. 2012). However, the inequality-enhancing impacts of rubber could affect the future viability of these programs. Indeed, previous research has suggested that distributive equity is of critical importance for local participants in voluntary programs like forest certification (Pinto and McDermott 2013). Thus, the design of any rubber sustainability initiative must consider contextual factors discussed previously, including expropriation of land and other natural resources driven by concessions.

Several studies have suggested that adverse consequences of concessions have outweighed their benefits (Ingalls et al. 2018, Hett et al. 2020, Friis et al. 2016, Kenney-Lazar 2012). Additionally, impact assessments of concessions remain very rare, and legal compliance is low (Hett et al. 2020). This underscores the importance of monitoring and evaluation and adherence to social and environmental safeguards in concessions development (see Table 4.1). Our results comport with previous research suggesting that smallholder rubber production systems have comparatively positive socioeconomic outcomes when juxtaposed with large-scale concessions (Kenney-Lazar and Wong 2016). Major rubber producing countries including Thailand, Vietnam, and India have all built smallholder support into rubber policies and institutions (Fox and Castella 2013). However, many voluntary sustainability programs have traditionally struggled with smallholder engagement (Pinto and McDermott 2012). Our findings suggest that this should be an area for continued focus in rubber sustainability programs.

The concern that voluntary programs for rubber could worsen existing economic inequality and polarization trajectories is also important to consider. Previous studies looking at community-based natural resource management have found that inequality significantly hampers participatory management efforts (Adhikari and Di Falco 2009). On a positive note, some research has suggested that forest certification improves benefits distribution, despite challenges with elite capture (Kalonga, Kulindwa, and Mshale 2015, Loveridge et al. 2021). Some programs have taken steps to enhance equity, for example through the smallholder certification fund of FSC and through revisions to its group and regional standards (Pinto and McDermott 2012, FSC 2019). Thus, voluntary sustainability programs – if deployed with appropriate safeguards – could help mitigate economic inequality and polarization impacts of rubber production.

4.6 Conclusion

In this analysis, we have explored changes in rural economic inequality and polarization in Laos from 2007/08 to 2012/2013 and their relationship to regional rubber production, based on rubber production area estimates derived from remotely sensed imagery and the LECS nationally representative household survey. Overall, our results suggest that rubber production is associated with worse inequality and polarization outcomes as compared to non-rubber areas in all three regions of the country. Further, our findings suggest that divergent rubber production systems in Laos are accompanied by different economic inequality and polarization trajectories.

This study contributes to the literature exploring the linkages between land-use change, agricultural commercialization, economic inequality, and polarization. Further, we bring attention to an understudied research area - the connections between economic inequality and voluntary sustainability programs.

With regards to economic inequality, we find that the income distribution becomes more equitable in the North and less so in the South. In the North, where smallholder production is dominant and rubber adoption began earlier, rubber is associated with increased per capita consumption and decreases in economic inequality that are slightly lower than those occurring in non-rubber areas. In the South, where rubber is primarily grown via more recent concessions, it is associated with larger increases in economic inequality. The Central region, home to a mix of individual farmers and concessions, exhibits higher increases in economic inequality in rubber production areas, as well as higher increases in per capita consumption. Our results suggest that smallholder rubber production systems may result in improved socioeconomic outcomes as compared to large-scale concessions. Revenues from smallholder production remain in the area, while those from concession-based production remain with investors, thus worsening rural economic inequality. Concessions are also more likely to drive land dispossession, reduced access to other natural resources, and generate income opportunities for local populations limited to poorly paid temporary wage labor – and all of these impacts are likely to have the greatest negative impacts on the poorest segments of the population (see Nanhthavong et al. 2022, *forthcoming*). Thus, our findings emphasize the need for improved regulation and enforcement for concessions (see Hett et al. 2020, Smith et al. 2020). Future research could compare these results to trends for other boom crops in Laos grown in individual, contract, and concession farming models. Moreover, our polarization analyses suggest that although rubber does not appear to be the largest exogenous driver of economic polarization overall (based on the Z-K index results), rubber production in the Southern and Central regions has been associated with increased polarization, as well as an increased number of households on the lower end of the income distribution in the South (as measured by the DER index).

Previous research indicates that specific aspects of rubber production systems may shape these inequality and polarization outcomes, including the extent of livelihood diversification, labor dynamics, prevalence of contract farming, and differential economic benefits from rubber production (Polthanee 2021, Evans et al. 2011, Baird and Fox 2015, Baird 2010, Yongou et al. 2019, Hett et al. 2020). Additional factors that may influence our results include baseline socioeconomic conditions (like ethnic group composition), large-scale agricultural commercialization trends, and other land concessions (Bader et al. 2017, Warr, Rasphone and Menon 2015, Hett et al. 2020, Ingalls et al. 2018, Schönweger et al. 2012).

Our analysis underscores the importance of incorporating equity-enhancing measures in the design of rubber sustainability interventions. Voluntary sustainability programs – in Laos and other Southeast Asian countries – should consider and engage the mechanisms by which economic inequality and polarization manifest in rubber production. This includes factors like tenure security and dispossession from land and forest resources, insufficient worker protections, potential livelihood vulnerability driven by price volatility, barriers for smallholders, and gender and ethnic inequality. If voluntary programs sufficiently address inequality in their design and implementation, they have the potential to help mitigate current trends. In fact, previous research has suggested that incentives for voluntary programs like FSC may help address sustainability challenges in the Laotian forest sector (Smith et al. 2021). However, voluntary measures alone are unlikely to shift the entire rubber sector into sustainability and must be complemented by government policy (see Kenney-Lazar et al. 2018). Finally, continued research is needed on the outcomes of voluntary sustainability programs for inequality - both for rubber and for other crops. Where positive examples of concessions do exist from a sustainability perspective, they should be highlighted (Smith et al. 2021).

The global and Laotian rubber markets have changed significantly since the rubber boom. Prices have decreased since 2011, forcing rubber producers to adapt to a shifting market (Vongvisouk and Dwyer 2016, Jin et al. 2021). The rubber moratorium is still in place and was reaffirmed in 2015, although the Government of Laos lifted suspensions on other plantation concessions in 2018 (Baird 2020, Smith et al. 2021). Some argue that the rubber moratorium has created a policy vacuum, as well as uncertainties for producers, incentivizing conversion of rubber plantations to other crops (Smith et al. 2020). However, studies project that rubber prices will continue to increase, with persistent market volatility (Norizan and Yusof 2021). Thus, coordinated efforts across the supply chain and among public and private stakeholders will continue to be necessary to reduce negative sustainability impacts of rubber production, including for rural economic inequality and polarization.

4.7 Bibliography

Adhikari, Bhim, and Salvatore Di Falco. 2009. "Social Inequality, Local Leadership and Collective Action: An Empirical Study of Forest Commons." The European Journal of Development Research 21 (2): 179–94. https://doi.org/10.1057/ejdr.2008.16.

Ahrends, A., Hollingsworth, P.M., Ziegler, A.D., Fox, J.M., Chen, H., Su, Y., Xu, J., 2015. Current trends of rubber plantation expansion may threaten biodiversity and livelihoods. Global Environmental Change 34, 48–58. https://doi.org/10.1016/j.gloenvcha.2015.06.002

Antonio Duro, Juan. 2010. "Decomposing International Polarization of per Capita CO2 Emissions." Energy Policy, Energy Efficiency Policies and Strategies with regular papers., 38 (11): 6529–33. https://doi.org/10.1016/j.enpol.2010.06.017.

Araar, Abdelkrim. 2006. "On the Decomposition of the Gini Coefficient: An Exact Approach, with an Illustration Using Cameroonian Data." SSRN Scholarly Paper ID 877147. Rochester, NY: Social Science Research Network. https://doi.org/10.2139/ssrn.877147.

Araar, Abdelkrim. 2008. "On the Decomposition of Polarization Indices: Illustrations with Chinese and Nigerian Household Surveys." SSRN Scholarly Paper ID 1136142. Rochester, NY: Social Science Research Network. https://doi.org/10.2139/ssrn.1136142.

Bader, Christoph, Sabin Bieri, Urs Wiesmann, and Andreas Heinimann. 2017. "Is Economic Growth Increasing Disparities? A Multidimensional Analysis of Poverty in the Lao PDR between 2003 and 2013." *The Journal of Development Studies* 53 (12): 2067–85. https://doi.org/10.1080/00220388.2016.1251587.

Baek, Jungho, and Guankerwon Gweisah. 2013. "Does Income Inequality Harm the Environment?: Empirical Evidence from the United States." Energy *Policy* 62 (November): 1434–37. https://doi.org/10.1016/j.enpol.2013.07.097.

Baird, Ian G. 2010. "Land, Rubber and People: Rapid Agrarian Changes and Responses in Southern Laos." Journal of Lao Studies, 47.

Baird, Ian G. 2014. "Degraded Forest, Degraded Land and the Development of Industrial Tree Plantations in Laos." *Singapore Journal of Tropical Geography* 35 (3): 328–44. https://doi.org/10.1111/sjtg.12076.

Baird, Ian G., William Noseworthy, Nghiem Phuong Tuyen, Le Thu Ha, and Jefferson Fox. 2019. "Land Grabs and Labour: Vietnamese Workers on Rubber Plantations in Southern Laos." Singapore Journal of Tropical Geography 40 (1): 50–70. https://doi.org/10.1111/sjtg.12261.

Basu, Kaushik. 2006. "Globalization, Poverty, and Inequality: What Is the Relationship? What Can Be Done?" *World Development*, The Impact of Globalization on the World's Poor, 34 (8): 1361–73. https://doi.org/10.1016/j.worlddev.2005.10.009.

Bennett, Elizabeth A. 2018. "Voluntary Sustainability Standards: A Squandered Opportunity to Improve Workers' Wages." *Sustainable Development* 26 (1): 65–82. https://doi.org/10.1002/sd.1691.

Castella, Jean-Christophe, Guillaume Lestrelin, Cornelia Hett, Jeremy Bourgoin, Yulia Rahma Fitriana, Andreas Heinimann, and Jean-Laurent Pfund. 2013. "Effects of Landscape Segregation on Livelihood Vulnerability: Moving From Extensive Shifting Cultivation to Rotational Agriculture and Natural Forests in Northern Laos." *Human Ecology* 41 (1): 63–76. https://doi.org/10.1007/s10745-012-9538-8.

Chancel, Lucas, Thomas Piketty, Emmanuel Saez, and Gabriel Zucman. 2022. "World Inequality Report 2022." World Inequality Lab. https://wir2022.wid.world/www-site/uploads/2021/12/WorldInequalityReport2022_Full_Report.pdf.

Chen, Chau-Nan, Tien-Wang Tsaur, and Tong-Shieng Rhai. 1982. "The Gini Coefficient and Negative Income." *Oxford Economic Papers* 34 (3): 473–78.

Committee for Planning and Cooperation, National Statistical Centre. 2004. "The Household of Lao PDR: Social and Economic Indicators. Lao Expenditure and Consumption Survey 2002/03. LECS 3." Vientiane, Lao PDR.

Correa-Parra, Juan, José Francisco Vergara-Perucich, and Carlos Aguirre-Nuñez. 2020. "Water Privatization and Inequality: Gini Coefficient for Water Resources in Chile." *Water* 12 (12): 3369. https://doi.org/10.3390/w12123369.

Cowell, Frank. 2000. "Measurement of Inequality." In Atkinson A. and Bourguignon F. (eds.) Handbook of Income Distribution. Amsterdam: Elsevier, p. 87-166.

Cramb, Rob, Vongpaphane Manivong, Jonathan C. Newby, Kem Sothorn, and Patrick S. Sibat. 2017. "Alternatives to Land Grabbing: Exploring Conditions for Smallholder Inclusion in Agricultural Commodity Chains in Southeast Asia." The Journal of Peasant Studies 44 (4): 939–67. https://doi.org/10.1080/03066150.2016.1242482.

de Bonafos, Hubert. 2020. "Review of Transparency And Traceability Tools And Solutions Prepared for the Global Platform for Sustainable Natural Rubber (GPSNR)." GPSNR-REP-A02-006. Global Platform for Sustainable Natural Rubber. https://sustainablenaturalrubber.org/reports/.

Dong, J., Xiao, X., Chen, B., Torbick, N., Jin, C., Zhang, G. and Biradar, C., 2013. Mapping deciduous rubber plantations through integration of PALSAR and multi-temporal Landsat imagery. Remote Sensing of Environment, 134, pp.392-402.

Douangsavanh, Linkham, Bansa Thammavong, and Andrew Noble. 2008. "Meeting Regional and Global Demands for Rubber: A Key to Poverty Alleviation in Lao PDR?" The Sustainable Mekong Research Network (Sumernet). http://lad.nafri.org.la/fulltext/2023-0.pdf

Douangsavanh, D.L., 2009. Rubber in the GMS: An Integrated Research Exercise on Rubber Development in Lao PDR. Sustainable Mekong Research Network.

Duclos, Jean-Yves, Joan Esteban, and Debraj Ray. 2004. "Polarization: Concepts, Measurement, Estimation." Econometrica 72 (6): 1737–72. https://doi.org/10.1111/j.1468-0262.2004.00552.x.

Ellis, Frank. 2000. "The Determinants of Rural Livelihood Diversification in Developing Countries." Journal of Agricultural Economics 51 (2): 289–302. https://doi.org/10.1111/j.1477-9552.2000.tb01229.x.

Epprecht, Michael, Anne-Kathrin Weber, Rasso Bernhard, Khamlouang Keoka, Thatheva Saphangthong, Vongpaphane Manivong, Phanxay Ingxay, et al. 2018. Atlas of Agriculture in the Lao PDR: Patterns and Trends between 1999 & 2011. Bern: Centre for Development and Environment (CDE), University of Bern, Switzerland, and Ministry of Agriculture and Forestry (MAF), Lao PDR, with Bern Open Publishing (BOP). https://boris.unibe.ch/116598/.

Epprecht, Michael, Nicolas Minot, Reno Dewina, Peter Messerli, and Andreas Heinimann. 2008. "The Geography of Poverty and Inequality in the Lao PDR." https://doi.org/10.7892/BORIS.34330.

Evans, Tom P., Khamla Phanvilay, Jefferson Fox, and John Vogler. 2011. "An Agent-Based Model of Agricultural Innovation, Land-Cover Change and Household Inequality: The Transition from Swidden Cultivation to Rubber Plantations in Laos PDR." *Journal of Land Use Science* 6 (2–3): 151–73. https://doi.org/10.1080/1747423X.2011.558602.

Fair Rubber Association. "Fair Rubber Standards Version 3.2." 2021. https://fairrubber.org/wp-content/uploads/2021/06/00-Fair-Rubber-Standards-VS-3.2.pdf.

Fox, Jefferson, Jean-Christophe Castella, and Alan D. Ziegler. 2014. "Swidden, Rubber and Carbon: Can REDD+ Work for People and the Environment in Montane Mainland Southeast Asia?" *Global Environmental Change* 29 (November): 318–26. https://doi.org/10.1016/j.gloenvcha.2013.05.011.

Friis, Cecilie, and Jonas Østergaard Nielsen. 2016. "Small-Scale Land Acquisitions, Large-Scale Implications: Exploring the Case of Chinese Banana Investments in Northern Laos." *Land Use Policy* 57 (November): 117–29. https://doi.org/10.1016/j.landusepol.2016.05.028.

Friis, Cecilie, Anette Reenberg, Andreas Heinimann, and Oliver Schönweger. 2016. "Changing Local Land Systems: Implications of a Chinese Rubber Plantation in Nambak District, Lao PDR." Singapore Journal of Tropical Geography 37 (1): 25–42. https://doi.org/10.1111/sjtg.12137.

FSC. "What Is FSC Doing for Smallholders?" 2019. Forest Stewardship Council. Accessed December 23, 2021. https://fsc.org/en/newsfeed/what-is-fsc-doing-for-smallholders.

FSC. "Natural Rubber." 2021. Forest Stewardship Council. Accessed December 13, 2021. https://fsc.org/en/for-businesses/natural-rubber.

Gitz V., Meybeck A., Pinizzotto S., Nair L., Penot E., Baral H., and Xu J. 2020. Sustainable Development of Rubber Plantations in a Context of Climate Change: Challenges and Opportunities. Center for International Forestry Research (CIFOR). https://doi.org/10.17528/cifor/007860.

"GPSNR Policy Framework." 2020. Global Platform for Sustainable Natural Rubber. https://rubberplatform.sharepoint.com/sites/GPSNRMembers/Shared%20Documents/Forms/AllItems.aspx?id=%2Fsites%2FGPSNRMembers%2FShared%20Documents%2FGPSNR%20Members%2F04%5FMembership%20Requirements%2FGPSNR%20Policy%20Framework%2Epdf&parent=%2Fsites%2FGPSNRMembers%2FShared%20Documents%2FGPSNR%20Members%2F04%5FMembership%20Requirements&p=true.

Global Platform on Sustainable Natural Rubber (GPSNR). "Global Platform for Sustainable Natural Rubber – A Multi Stakeholder Initiative." 2021. https://sustainablenaturalrubber.org/.

Halpern, Benjamin S., Carissa J. Klein, Christopher J. Brown, Maria Beger, Hedley S. Grantham, Sangeeta Mangubhai, Mary Ruckelshaus, et al. 2013. "Achieving the Triple Bottom Line in the Face of Inherent Trade-Offs among Social Equity, Economic Return, and Conservation." *Proceedings of the National Academy of Sciences* 110 (15): 6229–34. https://doi.org/10.1073/pnas.1217689110.

Hepp, Catherine M., Thilde Bech Bruun, and Andreas de Neergaard. 2019. "Transitioning towards Commercial Upland Agriculture: A Comparative Study in Northern Lao PDR." *NJAS: Wageningen Journal of Life Sciences* 88 (1): 57–65. https://doi.org/10.1016/j.njas.2018.11.001.

Hett, C., V. Nanhthavong, S. Hanephom, A. Phommachanh, B. Sidavong, K. Phouangphet, J. Lu, A. Shattuck, M. Ingalls, R. Bernhard, S. Phathitmixay, C. Phomphakdy, A. Heinimann, and M. Epprecht. 2020. Land Leases and Concessions in the Lao PDR: A characterization of investments in land and their impacts. Bern, Schweiz: Centre for Development and Environment (CDE), University of Bern.

Higman, Sophie, and Ruth Nussbaum. 2002. "How Standards Constrain the Certification of Small Forest Enterprises." ProForest.

Hurni, Kaspar, and Jefferson Fox. 2018. "The Expansion of Tree-Based Boom Crops in Mainland Southeast Asia: 2001 to 2014." *Journal of Land Use Science* 13 (1–2): 198–219. https://doi.org/10.1080/1747423X.2018.1499830.

Index Mundi. "Rubber - Daily Price - Commodity Prices - Price Charts, Data, and News - IndexMundi." Accessed December 17, 2021. https://www.indexmundi.com/commodities/?commodity=rubber&months=180.

Indra, Indra, Suahasil Nazara, Djoni Hartono, and Sudarno Sumarto. 2018. "Expenditure Inequality and Polarization in Indonesia, 2002-2012." *International Journal of Social Economics* 45 (10): 1469–86. https://doi.org/10.1108/IJSE-02-2017-0051.

Ingalls, Micah, Jean-Christophe Diepart, Nhu Truong, Daniel Hayward, Tony Neil, Chanthavone Phomphakdy, Rasso Bernhard, et al. 2018. State of Land in the Mekong Region. Bern, Switzerland and Vientiane, Lao PDR: Centre for Development and Environment, University of Bern, and Mekong Region Land Governance, with Bern Open Publishing. https://boris.unibe.ch/120285/.

International Rubber Study Group. 2014. "International Rubber Study Group (IRSG): Voluntary Sustainable Natural Rubber (SNR) Initiative's Criteria and Performance Indicators." http://snr-i.org/file/file/SNR-i_KPI_document_June_EN_for_Ref.pdf.

Jin, Shaoze, Shi Min, Jikun Huang, and Hermann Waibel. 2021. "Falling Price Induced Diversification Strategies and Rural Inequality: Evidence of Smallholder Rubber Farmers." World Development 146 (October): 105604. https://doi.org/10.1016/j.worlddev.2021.105604.

Junquera, V., Meyfroidt, P., Sun, Z., Latthachack, P., Grêt-Regamey, A., 2020. From global drivers to local land-use change: understanding the northern Laos rubber boom. Environmental Science & Policy 109, 103–115. https://doi.org/10.1016/j.envsci.2020.04.013

Junquera, Victoria, and Adrienne Grêt-Regamey. 2020. "Assessing Livelihood Vulnerability Using a Bayesian Network: A Case Study in Northern Laos." Ecology and Society 25 (4). https://doi.org/10.5751/ES-12049-250438.

Kalonga, Severin K., Kassim A. Kulindwa, and Baruani I. Mshale. 2015. "Equity in Distribution of Proceeds from Forest Products from Certified Community-Based Forest Management in Kilwa District, Tanzania." *Small-Scale Forestry* 14 (1): 73–89. https://doi.org/10.1007/s11842-014-9274-6.

Kenney-Lazar, Miles. 2012. "Plantation Rubber, Land Grabbing and Social-Property Transformation in Southern Laos." The Journal of Peasant Studies 39 (3–4): 1017–37. https://doi.org/10.1080/03066150.2012.674942.

Kenney-Lazar, Miles, and Grace Wong. 2016. "Challenges and Opportunities for Sustainable Rubber in the Lao People's Democratic Republic." Center for International Forestry Research (CIFOR). https://doi.org/10.17528/cifor/006244.

Kenney-Lazar, Miles, Grace Wong, Himlal Baral, and Aaron J. M. Russell. 2018. "Greening Rubber? Political Ecologies of Plantation Sustainability in Laos and Myanmar." *Geoforum* 92 (June): 96–105. https://doi.org/10.1016/j.geoforum.2018.03.008.

Keovilignavong, Oulavanh, and Diana Suhardiman. 2020. "Linking Land Tenure Security with Food Security: Unpacking Farm Households' Perceptions and Strategies in the Rural Uplands of Laos." Land Use Policy 90 (January): 104260. https://doi.org/10.1016/j.landusepol.2019.104260.

Kusakabe, Kyoko. 2015. "Gender Analysis of Economic Land Concessions in Cambodia and in Northern Laos: Case of Rubber Plantations." In Land grabbing, conflict and agrarian-environmental transformations: perspectives from East and Southeast Asia, 19. Chiang Mai University. https://www.eur.nl/sites/corporate/files/CMCP_43-Kusakabe.pdf

Kusakabe, Kyoko, and Chanthavisith Chanthoumphone. 2021. "Transition From Subsistence Agriculture to Rubber Plantations in Northern Laos: Analysis of Household Livelihood Strategies by Ethnicity and Gender." SAGE Open 11 (2): 21582440211011464. https://doi.org/10.1177/21582440211011463.

Lao Department of Statistics, Ministry of Planning and Investment. 2010. "Lao Expenditure and Consumption Survey 2007/08: Study Documentation." Lao People's Democratic Republic.

Lao People's Democratic Republic: Department of Statistics, Ministry of Planning and Investment. 2010. "Lao Expenditure and Consumption Survey 2007/08: Study Documentation."

Law, Elizabeth A., Nathan J. Bennett, Christopher D. Ives, Rachel Friedman, Katrina J. Davis, Carla Archibald, and Kerrie A. Wilson. 2018. "Equity Trade-Offs in Conservation Decision Making." Conservation Biology 32 (2): 294–303. https://doi.org/10.1111/cobi.13008.

Loveridge, Robin, Susannah M Sallu, Marion Pfeifer, Johan A Oldekop, Mercy Mgaya, Daniel A da Silva, Julia Latham, Philip J Platts, and Andrew R Marshall. 2021. "Certified Community Forests Positively Impact Human Wellbeing and Conservation Effectiveness and Improve the Performance of Nearby National Protected Areas." Conservation Letters: e12831. https://doi.org/10.1111/conl.12831.

Lu, Juliet N. 2017. "Tapping into Rubber: China's Opium Replacement Program and Rubber Production in Laos." The Journal *of Peasant Studies* 44 (4): 726–47. https://doi.org/10.1080/03066150.2017.1314268.

Malkamäki, Arttu, Dalia D'Amato, Nicholas J. Hogarth, Markku Kanninen, Romain Pirard, Anne Toppinen, and Wen Zhou. 2018. "A Systematic Review of the Socio-Economic Impacts of Large-Scale Tree Plantations, Worldwide." *Global Environmental Change* 53 (November): 90–103. https://doi.org/10.1016/j.gloenvcha.2018.09.001.

Marion Karmann and Alan Smith. 2009. "FSC Reflected in Scientific and Professional Literature: Literature Study on the Outcomes and Impacts of FSC." Forest Stewardship Council. https://www.fsc-deutschland.de/preview.fsc-impact-study-engl.a-1341.pdf

Messerli, Peter, Andreas Heinimann, Michael Epprecht, Phonesaly Souksavath, Thiraka Chanthalanouvong, and Nicholas Minot. 2008. *Socio-Economic Atlas of the Lao PDR. An Analysis Based on the 2005 Population and Housing Census*. Bern: Geographica Bernensia. http://www.laoatlas.net/links/PDF%20download.html.

Michelin – DOA. "Responsible Management of the Natural Rubber Supply Chain." 2021. Accessed November 10, 2021. https://purchasing.michelin.com/en/responsible-managment-natural-rubber-supply-chain/

Milanovic, Branko. 2016. "Income Inequality Is Cyclical." Nature 537 (7621): 479–82. https://doi.org/10.1038/537479a.

Ministry of Agriculture and Forestry, Government of the Lao People's Democratic Republic. (MAF). 2014. "Lao Census of Agriculture 2010/11 Analysis of Selected Themes." https://www.fao.org/3/at767e/at767e.pdf.

Mookherjee, Dilip, and Anthony F. Shorrocks. 1982. "A Decomposition Analysis of the Trend in UK Income Inequality." Economic Journal 92 (368): 886–902.

Nanhthavong, Vong, Michael Epprecht, Cornelia Hett, Julie G. Zaehringer, and Peter Messerli. 2020. "Poverty Trends in Villages Affected by Land-Based Investments in Rural Laos." Applied Geography 124 (November): 102298. https://doi.org/10.1016/j.apgeog.2020.102298.

Nanhthavong, Vong; Nguyen, Anh-Thu; Hett, Cornelia; Bieri, Sabin; Epprecht, Michael. Forthcoming, 2022. Proletarianization and gateways to precarization in the context of land-based investments for agricultural commercialization in Lao PDR, World Development.

Norizan, Nor Farah Hanim Binti Mohamad, and Zahayu Binti Md Yusof. 2021. "Forecasting Natural Rubber Price in Malaysia by 2030." Malaysian Journal of Social Sciences and Humanities (MJSSH) 6 (9): 382–90. https://doi.org/10.47405/mjssh.v6i9.986.

Open Development Mekong. 2021. "Laos Administrative Boundaries (Country, Province and District) - OD Mekong Datahub." https://data.opendevelopmentcambodia.net/en/dataset/laos-administrative-boundaries-country-province-and-district.

Otten, Fenna, Jonas Hein, Hannah Bondy, and Heiko Faust. 2020. "Deconstructing Sustainable Rubber Production: Contesting Narratives in Rural Sumatra." *Journal of Land Use Science* 15 (2–3): 306–26. https://doi.org/10.1080/1747423X.2019.1709225.

Özdoğan, Mutlu, Ian G. Baird, and Michael B. Dwyer. 2018. "The Role of Remote Sensing for Understanding Large-Scale Rubber Concession Expansion in Southern Laos." *Land* 7 (2): 55. https://doi.org/10.3390/land7020055.

PEFC. "PEFC Supporting Sustainable Rubber: Official Campaign Launch." 2021. https://www.pefc.org/events-training/pefc-supporting-sustainable-rubber-official-campaign-launch.

Phimmavong, S., B. Ozarska, S. Midgley, and R. Keenan. 2009. "Forest and Plantation Development in Laos: History, Development and Impact for Rural Communities." *International Forestry Review* 11 (4): 501–13. https://doi.org/10.1505/ifor.11.4.501.

Phimmavong, Somvang, and Rodney J. Keenan. 2020. "Forest Plantation Development, Poverty, and Inequality in Laos: A Dynamic CGE Microsimulation Analysis." *Forest Policy and Economics* 111 (February): 102055. https://doi.org/10.1016/j.forpol.2019.102055.

Pinto, Luís Fernando Guedes, and Constance McDermott. 2012. "Equity and Forest Certification – Issues for debate from a case study in Brazil." Imaflora.

Pinto, Luís Fernando Guedes, and Constance McDermott. 2013. "Equity and Forest Certification — A Case Study in Brazil." Forest Policy and Economics 30 (May): 23–29. https://doi.org/10.1016/j.forpol.2013.03.002.

Pirelli. 2021. "Sustainable Natural Rubber Policy." http://www.pirelli.com/corporate/en-ww/governance/principles/group-policy/gov-sustainable-natural-rubber-policy.

Plat, Didier. 2012. Inequality and Concentration Indices and Curves. R Package Version 1.0-1. R. https://CRAN.R-project.org/package=IC2.

Polthanee, Anan, Arunee Promkhambut, Satit Aditto, and Avakat Phasouysaingam. 2021. "Impact of Changing Swidden-Based Farming to Rubber-Based Farming on Food

Security in Luang Namtha Province, Lao PDR." Sustainability 13 (14): 7617. https://doi.org/10.3390/su13147617.
Potapov, P., A. Tyukavina, S. Turubanova, Y. Talero, A. Hernandez-Serna, M. C. Hansen, D. Saah, et al. 2019. "Annual Continuous Fields of Woody Vegetation Structure in the Lower Mekong Region from 2000-2017 Landsat Time-Series." Remote Sensing of Environment 232 (October): 111278. https://doi.org/10.1016/j.rse.2019.111278.
Potapov, Peter, Matthew C. Hansen, Indrani Kommareddy, Anil Kommareddy, Svetlana Turubanova, Amy Pickens, Bernard Adusei, Alexandra Tyukavina, and Qing Ying. 2020. "Landsat Analysis Ready Data for Global Land Cover and Land Cover Change Mapping." *Remote Sensing* 12 (3): 426. https://doi.org/10.3390/rs12030426.
Qaim, Matin, Kibrom T. Sibhatu, Hermanto Siregar, and Ingo Grass. 2020. "Environmental, Economic, and Social Consequences of the Oil Palm Boom." Annual *Review of Resource Economics* 12 (1): 321–44. https://doi.org/10.1146/annurev-resource-110119-024922.
QGIS.org, 2021. QGIS Geographic Information System. QGIS Association. http://www.qgis.org
R Core Team. 2013. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-project.org/
Ravallion, Martin. 2005. "A Poverty-Inequality Trade Off?" *The Journal of Economic Inequality* 3 (2): 169–81. https://doi.org/10.1007/s10888-005-0091-1.
Reardon, Thomas, J. Edward Taylor, Kostas Stamoulis, Peter Lanjouw, and Arsenio Balisacan. 2000. "Effects of Non-Farm Employment on Rural Income Inequality in Developing Countries: An Investment Perspective." *Journal of Agricultural Economics* 51 (2): 266–88. https://doi.org/10.1111/j.1477-9552.2000.tb01228.x.
Ricci, Chiara Assunta, and Sergio Scicchitano. 2021. "Decomposing Changes in Income Polarization by Population Group: What Happened during the Crisis?" *Economia Politica* 38 (1): 235–59. https://doi.org/10.1007/s40888-021-00218-8.
Rigg, Jonathan. 2005. "Poverty and Livelihoods after Full-Time Farming: A South-East Asian View." *Asia Pacific Viewpoint* 46 (2): 173–84. https://doi.org/10.1111/j.1467-8373.2005.00266.x.
Schönweger Oliver, Heinimann Andreas, Epprecht Michael, Lu Julliet, and Thalongsengchanh P. 2012. *Concessions and Leases in the Lao PDR: Taking Stock of Land Investments*. Geographica Bernensia. https://doi.org/10.13140/RG.2.1.3501.7047.
Schulenberg, René (2018). dineq: Decomposition of (Income) Inequality. R package version 0.1.0. https://CRAN.R-project.org/package=dineq
Scruggs, Lyle A. 1998. "Political and Economic Inequality and the Environment." *Ecological Economics* 26 (3): 259–75. https://doi.org/10.1016/S0921-8009(97)00118-3.
Silber, Jacques. 2012. *Handbook of Income Inequality Measurement*. Springer Science & Business Media.
Silva, Julie A., Corene J. Matyas, and Benedito Cunguara. 2015. "Regional Inequality and Polarization in the Context of Concurrent Extreme Weather and Economic Shocks." *Applied Geography*, Spatial Inequality, 61 (July): 105–16. https://doi.org/10.1016/j.apgeog.2015.01.015.
Singh, Sarinda. 2020. "Understanding Changes for Forest-Dependent Communities in Upland Laos." *Human Ecology* 48 (2): 199–209. https://doi.org/10.1007/s10745-020-00143-7.
Sithong, Thongmanivong, and Vongvisouk Thoumthone. 2006. "Impacts of cash crops on rural livelihoods: a case study from Muang Sing, Luang Namtha Province, Northern Lao

PDR." In *Hanging in the Balance: Equity in Community-Based Natural Resource Management in Asia,* edited by Sango Mahanty, Jefferson Fox, Michael Nurse, Peter Stephen, Leslie McLees. RECOTC and East-West Center: 106-121.

Smith, H., Lu, J., To, P.X., Mienmany, S., Soukphaxay, K., 2020. Rubber Plantation Value Chains in Laos: Opportunities and Constraints in Policy, Legality and Wood Processing. Report produced for ACIAR project FST/2016/151 - Advancing enhanced wood manufacturing industries in Laos And Australia, and Forest Trends. https://www.forest-trends.org/wp-content/uploads/2020/07/Rubber-Plantation-Value-Chains-in-Laos.pdf

Smith, Hilary, Peter Kanowski, Rodney J. Keenan, and Somvang Phimmavong. 2021. "Lao Plantation Policy: Prospects for Change." *Forests* 12 (8): 1132. https://doi.org/10.3390/f12081132.

Suhardiman, Diana, Mark Giordano, Oulavanh Keovilignavong, and Touleelor Sotoukee. 2015. "Revealing the Hidden Effects of Land Grabbing through Better Understanding of Farmers' Strategies in Dealing with Land Loss." Land Use Policy 49 (December): 195–202. https://doi.org/10.1016/j.landusepol.2015.08.014.

Pimhidzai, Obert, Nina Fenton, Phonesaly Souksavath, and Vilaysouk Sisoulath. 2014. "Poverty Profile In Lao PDR: Poverty Report For The Lao Consumption And Expenditure Survey 2012-2013." Ministry of Planning and Investment, Lao Statistics Bureau. https://openknowledge.worldbank.org/bitstream/handle/10986/23023/Poverty0profil0ure0survey0201202013.pdf;jsessionid=FC5C8CBDB325369ED5546395F5963D41?sequence=1.

Vongvisouk, Thoumthone, and Michael Dwyer. 2016. "Falling Rubber Prices in Northern Laos: Local Responses and Policy Options." Helvetas.

Vongvisouk, Thoumthone, Ole Mertz, Sithong Thongmanivong, Andreas Heinimann, and Khamla Phanvilay. 2014. "Shifting Cultivation Stability and Change: Contrasting Pathways of Land Use and Livelihood Change in Laos." Applied Geography 46 (January): 1–10. https://doi.org/10.1016/j.apgeog.2013.10.006.

Warr, Peter, Sitthiroth Rasphone, and Jayant Menon. 2015. *Two Decades of Rising Inequality and Declining Poverty in the Lao People's Democratic Republic*. Asian Development Bank. https://www.adb.org/publications/two-decades-rising-inequality-declining-poverty-lao-pdr.

Warren-Thomas, Eleanor, Paul M. Dolman, and David P. Edwards. 2015. "Increasing Demand for Natural Rubber Necessitates a Robust Sustainability Initiative to Mitigate Impacts on Tropical Biodiversity." *Conservation Letters* 8 (4): 230–41. https://doi.org/10.1111/conl.12170.

Willer, Helga, Gregory Sampson, Vivek Voora, Duc Dang, and Julia Lernoud. 2019. State of Sustainable Markets 2019: Statistics and Emerging Trends. Place of publication not identified: International Trade Center. https://www.deslibris.ca/ID/10102592.World Food Programme (WFP). "Food and Nutrition Security Atlas of Lao PDR." 2013. https://documents.wfp.org/stellent/groups/public/documents/ena/wfp260762.pdf.

Yongnou, Xonglao, Sithong Thongmanivong, Kaisone Phengsopha, and Thoumthone Vongvisouk. 2019. "Impacts of Land Concession on Rural Livelihoods in Central Laos." International Journal of Science, Environment and Technology 8 (4): 735–46.

Zhang, X., and R. Kanbur. 2001. "What Difference Do Polarisation Measures Make? An Application to China." The Journal of Development Studies 37 (3): 85–98. https://doi.org/10.1080/00220380412331321981.

Chapter 5: Conclusion

5.1 Introduction

Despite persistent gaps in the science on the effectiveness of VSPs, there is burgeoning interest in this topic sparked by recent environmental commitments, including deforestation and other land-based emissions reduction pledges arising from the 2021 United Nations' Conference of the Parties (Chandrasekhar and Viglione 2021). Thus, the empirical findings presented in this dissertation grapple with a central question in sustainability research: how can we best measure the impact of voluntary sustainability programs (VSPs)? I explored this topic at a global scale through Chapter 2, as well as two national case studies of Lao PDR (Laos) in Chapters 3 and 4. Laos is used as case study to better understand broader dynamics relevant for other countries where VSPs are increasingly implemented.

Through each chapter, I investigated three interconnected questions related to VSP impacts: (1) which programs, geographies, and outcomes have been studied utilizing counterfactual approaches, and what are the research gaps, (2) to what extent do a specific set of VSPs impact forest cover and poverty, based on a case study of Laos from 2005-2015, and (3) to what degree should voluntary programs engage with and address economic inequality – an under-studied sustainability outcome – in the context of rapid agricultural intensification? I employed a diverse set of methods to explore these issues, including a systematic literature review, inequality and polarization analysis, policy and content analysis, and causal inference via statistical matching. I concentrated on agriculture and forestry VSPs due to the importance of these sectors for sustainable development, including for poverty alleviation, ecosystem health, and climate change adaptation.

This dissertation makes several key contributions to sustainability research, policy, and practice. I synthesize the existing body of impact evaluation literature on a specific type of voluntary intervention (agricultural voluntary sustainability standards or VSSs), clearly identifying and outlining future research needs. I also conduct a novel assessment of the outcomes of sustainability programs in Laos using directed acyclic graphs (DAGs) to guide statistical matching, providing the foundation for future counterfactual analyses of environmental programs. Finally, I highlight the importance of economic inequality as a sustainability outcome in the design and evaluation of VSPs, informing their future implementation.

5.2 Summary of Major Findings

5.2.1 Chapter 1: Introduction

Chapter 1 provides an overview of VSPs, with a particular focus on those included in this study – VSSs and SDPs. It describes important aspects of VSP structure including the provision of resources and incentives to participants to offset opportunity costs, the potential for sustainability trade-offs in program implementation, and distinctions between VSSs and SDPs. Chapter 1 also underscores the importance of counterfactual evaluations of VSPs, which eliminate rival explanations for program outcomes through their research design and methodology.

5.2.2 Chapter 2: Progress and Pitfalls: A systematic review of the evidence for agricultural sustainability standards

VSSs are a key approach leveraged by NGOs and the private sector to foster sustainable commodity production; certified lands are estimated to cover about 1.1% of global agricultural area (Tayleur et al. 2017). However, counterfactual evidence for VSSs is notoriously lacking.

Chapter 2 surveys 45 studies using counterfactual approaches for 13 major agricultural VSSs, as well as spatial and statistical data on the global extent of VSS certification. This chapter addresses three key questions: (1) which VSSs are being studied and where (and how does this compare to the extent of certification), (2) which outcomes are included in evaluations, and (3) what does the current evidence base suggest regarding outcomes of agricultural VSSs? For the first research question, I identify a major imbalance in the coverage of impact evaluations of agricultural VSSs. Among studies using quasi-experimental approaches (e.g., propensity score matching, instrumental variables, difference-in-difference analyses), coffee is by far the most studied crop, with 75% of the cases analyzed. I also find that economic sustainability is studied most frequently, and only 20% of studies assess all three pillars of sustainability—economic, social, and environmental—simultaneously. The narrow focus on economic outcomes constrains our understanding of trade-offs and co-benefits in VSS implementation. No identified studies used quasi-experimental approaches to assess economic inequality impacts, a gap that is explored further in Chapter 4.

The findings of Chapter 2 indicate that the economic and environmental outcomes of VSSs tend to be positive; however, social outcomes (e.g. local satisfaction, gender equality, strengthening social networks) typically exhibit no difference between certified and noncertified producers. Finally, I describe contextual factors influencing certification’s outcomes based on the literature. For example, existing poverty and market conditions shape whether certification is the appropriate sustainability intervention; if poverty is severe, other interventions may be more effective.

5.2.3 Chapter 3: The poverty and forest cover impacts of sustainable development projects in Lao PDR from 2005 to 2015

Based on the methods reviewed in Chapter 2, I conducted a nationally representative counterfactual analysis of SDPs in Laos from 2005-2015, focusing on forest cover and poverty. Since effect heterogeneity is a known occurrence in VSP implementation, I assessed the average treatment effect on the treated (ATT) for four VSP groups – (1) all projects together, (2) food security and livelihoods projects, (3) forest management projects, and (4) sustainable resource management projects. I utilized inverse probability of treatment weighting (IPTW) to adjust for baseline differences among treatment and control villages prior to conducting a weighted regression.

The results suggest that VSPs in Laos had an overall positive effect during the study period. However, there was some effect heterogeneity among different project types. For example, food security and livelihoods projects exhibited significant, positive impacts for both poverty and proportional forest cover in 2015 (3.7% and .9 percentage points/pp respectively[21]). Sustainable resource management projects showed the highest positive effects for forest cover (1.8 pp), but no significant effects for poverty. Sustainable forest management projects showed no significant effects for either outcome.

The results for food security and livelihoods projects suggest strong potential for co-benefits, which comports with previous research on the drivers of poverty within Laos, as well as SDP project evaluations. There are several potential explanations for the inconclusive forest management project results. For example, for poverty outcomes, challenges hampering SDP success may have included foregone income for resource extraction, pre-existing land and

[21] % changes refer to relative change; percentage point or pp changes refer to absolute change

income inequalities, land tenure insecurity, market conditions, limited producer organizations, and/or insufficient participant incentives. For forest cover, forest management SDPs may not have addressed key drivers of forest loss (e.g. illegal logging, large-scale concessions). Overall, though, this chapter tells a promising story about the impact of SDPs on forest cover, since three out of four models suggest favorable outcomes.

5.2.4 Chapter 4: Cultivating Inequality? Regional rubber dynamics and implications for voluntary sustainability programs in Lao PDR

Chapter 4 examines two understudied topics in the evaluation of VSPs - economic inequality and polarization. Specifically, it assesses changes in rural economic inequality and polarization in the latter period of the Laos rubber boom (2007/08-2012/13) through the calculation and decomposition of the Gini and Duclos Esteban Ray indices. The analysis combines household level per capita consumption data and rubber production area estimates. I evaluate trends in Northern, Central, and Southern Laos separately, which all represent different production system compositions. Northern production is primarily comprised of smallholder systems, the Central region represents a mix of smallholders, contract farming, and concessions, and Southern Laos is dominated by large-scale concessions. The findings suggest that overall, rubber production is associated with increased economic inequality and polarization, but the magnitude of this relationship differs based on the regional production system in Laos.

I hypothesized that Southern Laos would exhibit negative economic inequality and polarization associations as compared to Northern and Central Laos, due to significant literature outlining the deleterious impacts of the concession model on local livelihoods in that region (Baird and Fox 2015, Baird 2010, Smith et al. 2020, Fox and Castella 2013). Problems with concessions include low wages and seasonal piece rate pay for local laborers, poor working conditions (e.g. lack of contracts), and land and resource dispossession (exacerbated by comparatively insecure land tenure in the South). The findings uphold this relationship, with striking increases in economic inequality and polarization in Southern rubber production areas. Further, the findings illustrate that smallholder systems in Northern Laos were associated with the "least bad" economic inequality changes – rural consumption became more equal for rubber and nonrubber producing groups, and the Northern rubber group exhibited increased expenditures. The discussion outlines potential reasons for this, including the development of smallholder rubber systems in the North via social and investor networks, and potential co-benefits of smallholder rubber production like livelihood diversification and increased land tenure security.

This analysis suggests that the concession model has serious limitations for promoting equitable economic outcomes for farmers, necessitating stronger regulation and enforcement regarding land tenure security and worker protections. Further, since concession wage laborers tend to be poor, the inequality-enhancing impacts of concessions impose disproportionately high costs on disadvantaged groups. Overall, the results underline the importance of economic inequality and polarization for VSPs, particularly those operating in contexts of recent or ongoing crop booms, or where concessions dominate.

5.3 Empirical contributions

The research presented in this dissertation makes three key contributions to a better understanding of VSP impacts, namely: (1) innovative data synthesis to explore interlinked socioeconomic and environmental systems, (2) contributions to the field of causal inference and environmental interventions through rigorous methodological approaches, and (3) provision of expansive policy insights for sustainability interventions.

First, this research represents an inventive synthesis of LCLUC data (e.g. rubber production area, forest cover) with socioeconomic data (expenditures, poverty) to investigate policy-relevant questions about VSP effectiveness. For Chapter 4, this involved navigating the different temporal and geographical scales of socioeconomic and LCLUC data. For Chapter 3, this involved the aggregation of 32 socioeconomic and biophysical covariates to village land-use area polygons in Laos, a significant data processing undertaking. Additionally, across the entire dissertation, the underlying analysis materials (R scripts, data, etc.) were or will be made publicly available upon publication, enabling future researchers to build upon or replicate the results. For Chapter 2, this included a publicly available database of all indicators and studies I incorporated in the systematic review of agricultural VSSs. For Chapters 3 and 4, I will provide the analysis code and pre-processed, streamlined versions of the analysis datasets upon request to enable replication (with appropriate re-labeling and anonymization of households and villages).

Further, this dissertation makes a novel contribution to the literature on causal inference of VSP impact. Chapter 3 represents, to my knowledge, the first study utilizing directed acyclic graphs to design an inverse probability-of-treatment weighted regression (IPWR) analysis for sustainability programs. DAGs are heavily utilized in other fields to minimize bias in counterfactual research. In addition to its methodological and empirical contributions, the application of DAGs to VSP evaluation is useful for practitioners, and showcases a unique example of transdisciplinary sustainability science.

Finally, from a policy perspective, this dissertation's empirical findings can be applied to inform and improve all aspects of the VSP project cycle. For example, Chapter 4's findings that the rubber boom increased economic inequality – particularly in large-scale concessions – provide critical context for rubber and forestry-focused VSPs, and enable robust discussion about how the design, implementation, and evaluation of these programs should shift to better account for economic inequality. The treatment effect estimates in Chapter 3 not only provide evidence that sustainable development-focused VSPs can successfully drive positive forest cover outcomes – they also help illustrate the magnitude of positive impacts that can occur, informing strategy development and target-setting for stakeholders. Chapter 3's empirical findings regarding co-benefits and trade-offs can also help inform future VSP design, for example related to the provision of incentives and training to participants, and the importance of addressing both resource management and livelihoods objectives concurrently.

5.4 Policy relevance and contributions

Each chapter is informed by policy-relevant questions and puts forth policy recommendations based on research findings. These recommendations are useful for a broad set of stakeholders, including NGOs, academics, the private sector, governments, and local communities. Taken together, the research results hold four key implications for policy: (1) the evidence base for VSPs is not representative of participating regions and sectors; (2) different types of VSPs bring different sustainability benefits and trade-offs; (3) overall, VSPs can be successful in driving positive poverty and forest cover outcomes; (4) it is critical to monitor and address the impact of VSPs on economic inequality.

First, Chapter 2's findings suggest that further impact evaluation research is needed for understudied crops, certifications, and countries for which there is significant certified agricultural area. This is particularly true in cases where agricultural production is a known driver of serious environmental problems. Palm oil, soy, cocoa, sugar cane, and cotton VSSs had all seen little to no counterfactual research at the time of my review. Since then, some additional research has emerged evaluating palm oil and cocoa, which hopefully will engender further inquiry (Santika et al. 2020, Sellare et al. 2020, Sellare, Meemken, and Qaim 2020). Geographical evidence gaps included Brazil, Malaysia, the Ivory Coast, and the United States. There is a need to address this through multi-stakeholder applied research (with collaboration from NGOs, academics, supply chain stakeholders, etc.). Soon after publication of this chapter, I provided feedback into the new iteration of World Wildlife Fund's Certification Assessment Tool, as well as the organization's broader approach to sustainability standards, illustrating the tangible policy value of this work.

Second, Chapter 3's findings suggest that some VSPs (food security and livelihoods projects) are more likely to drive positive outcomes for both poverty and forest cover, while other VSPs (resource management projects) may generate positive impacts for forest cover, but not for poverty. These findings illustrate the importance of prioritizing alternative livelihoods in community-level VSP implementation, e.g. via agroforestry. The results also underscore the need for future research regarding impacts of forest management projects, and highlight several challenges they encounter related to local socioeconomic and market conditions. Generally, though, Chapter 3 demonstrates that positive forest cover outcomes are possible for voluntary programs, thus providing empirical evidence that these approaches should remain a component of our broader strategy to address environmental problems.

Chapter 4's findings suggest that economic inequality is important to consider and address in VSPs; however, inequality outcomes are currently understudied in the literature. Additionally, rubber VSPs – like many others – have neglected to monitor economic inequality as a challenge to environmental sustainability. I put forth a set of recommendations for VSPs to address economic inequality in rubber production systems for workers, local communities, and smallholder farmers. These include enforcing workers' protection and rights, supporting collective organizing for small farmer associations, and addressing potential livelihood vulnerability caused by commodity market fluctuations (e.g. via livelihood diversification or support for floor prices).

Although this analysis is specific to the rapid increase of natural rubber in Laos, the recommendations are relevant in other cases of rapid agricultural intensification, where large land concessions are employing local labor and/or smallholders face barriers to entry. The analysis also highlights significant limitations of Laos' rubber concession model in terms of socioeconomic inequality. At the time of journal submission, preliminary findings were shared with select practitioners involved in the Global Platform for Sustainable Natural Rubber (GPSNR), which is currently refining its reporting requirements based on stakeholder negotiations. Practitioners shared that the analysis clearly illustrated the risk of negative local economic impacts from rubber concessions (for example due to low wages for laborers and resulting inequality) - a risk which private sector entities often do not wish to discuss in multi-stakeholder programs. Thus, the analysis will help advocate for the importance of strong worker protection mechanisms in the GPSNR.

5.5 Recommendations for future work

Based on my experiences conducting the research described in this dissertation, I set forth three main recommendations for future work: (1) incorporate primary data collection and analysis evaluating socioeconomic and environmental dynamics in VSP research, (2) enhance the use of Earth observation data in analysis of VSP outcomes, and (3) diversify and strengthen the evidence base for VSP impact.

As originally conceived, field research was an integral component of this dissertation. However, this study was conducted from 2019 to 2022, and field data collection was impossible due to the COVID-19 pandemic. In the future, field-collected data could help contextualize and/or supplement the research and assess the generalizability of the study findings. For example, building on Chapter 4, field data could illuminate local experiences regarding economic inequality and natural rubber production in Laos. Interviews or surveys could also evaluate other types of inequality transitions associated with rubber production, for example related to social dynamics, gender, or ethnic minority group status. In the case of Chapter 3, field research could provide data regarding local perceptions of VSP effectiveness. Other topics for future research posed by these chapters include the impact of rapid rubber price drops on local livelihoods since 2013, the local impacts of forest management VSPs, and the impact of VSPs on economic inequality.

Beyond the general need for interviews and surveys to provide more context on specific topics, there are a few ways in which field insights could provide particular value for continued work on Chapter 3. One potential strategy could involve a sample of VSP villages (selected, for example, to reflect a variety of the different SDP project types), and include interview–surveys and focus groups of community members, participating NGOs, and potentially other stakeholders, like extension officers and/or government officials. Field data collection could assist in the following: (1) forming a more detailed impression of the implementation and design of a subset of the VSPs, particularly the extent to which participatory approaches were employed, (2) identifying which sustainability outcomes were most important to the participants, and the extent to which participants experienced changes in those outcomes, (3) ascertaining which mechanisms were key in driving sustainability outcomes as experienced by the participants (e.g. income from alternative livelihoods versus infrastructure improvements), and (4) identifying any relevant trade-offs experienced by the participants. This would enable an exploration of the project implementation and design factors that may have influenced sustainability outcomes. Ideally, fieldwork would also take a three – pillar approach, incorporating consideration of at least one social and one economic outcome, as well as finer-scale consideration of the Landsat forest cover data.

Additionally, there is a distinctive role in VSP research for the application of remotely sensed LCLUC data to measure environmental outcomes. This should be complemented by direct measurement of environmental changes due to VSPs (e.g. forest changes and carbon sequestration). This approach could be leveraged for many different types of programs, including the understudied VSSs identified in Chapter 2. Some of the highest priority VSSs to evaluate include the Roundtable on Sustainable Palm Oil, the Better Cotton Initiative, Bonsucro, and the Roundtable on Responsible Soy, as well as cocoa certified under Rainforest Alliance and other programs.

There are also several ways in which the VSP evidence base should be further developed and strengthened. For all VSPs, research assessing socioeconomic and environmental outcomes simultaneously holds a unique value and should be prioritized. Further, VSPs which have seen less evaluation in the quasi-experimental literature – including zero deforestation commitments, VSSs, REDD+, and 'nature-based solutions' as a whole – should receive particular attention. Finally, there is a broader need to clearly outline causal assumptions *and* enable replication in studies evaluating VSP impact.

5.6 Concluding remarks

This research examines the evaluation of VSPs' economic, social, and environmental outcomes, both globally (Chapter 2) and using Laos as a case study (Chapters 3 and 4). I find that while the evidence base for VSP's is uneven, previous empirical work suggests that voluntary sustainability standards for agriculture often generate economic and environmental benefits. However, social outcomes are less consistently positive for VSS's, indicating an area for continued improvement and adaptive management. Then, using a counterfactual approach, I demonstrate that agriculture and forestry SDPs have generated some positive and simultaneous outcomes – both for socioeconomic sustainability, e.g. via poverty reduction, and for environmental sustainability, as measured by forest cover. This evidence lends support to the continued utilization of VSPs as an environmental intervention, underpinned by regulatory approaches. Importantly, findings indicate that in Laos, programs focused on food security and livelihoods also deliver environmental co-benefits. Finally, I find that the inequality-enhancing aspects of rubber production, particularly under the large-scale concession model, could pose additional challenges for voluntary programs. Thus, evaluations of inequality should be incorporated into future VSP assessments.

While growing, the VSP evidence base is still inhibited by research imbalances (some intervention types and geographies are studied more than others), and the fact that outcomes are highly context dependent and shaped by external drivers. In Laos in particular, I theorize that external drivers related to land tenure insecurity, concession development, and illegal logging limit the effectiveness of forest management VSPs. Further, VSPs must navigate the (sometimes negative) surrounding sustainability changes resulting from market transitions, including agricultural intensification and crop booms. My findings highlight the need to investigate potential complementarity between voluntary and regulatory approaches and different types of voluntary interventions for varied sustainability outcomes, particularly through counterfactual research designs.

5.7 Bibliography

Baird, Ian G. 2010. "Land, Rubber and People: Rapid Agrarian Changes and Responses in Southern Laos." Journal of Lao Studies, 47.

Baird, Ian G., and Jefferson Fox. 2015. "How Land Concessions Affect Places Elsewhere: Telecoupling, Political Ecology, and Large-Scale Plantations in Southern Laos and Northeastern Cambodia." Land4 (2): 436–53. https://doi.org/10.3390/land4020436.

Chandrasekhar, Aruna, and Giuliana Viglione. 2021. "COP26: Key Outcomes for Food, Forests, Land Use and Nature in Glasgow." Carbon Brief. November 17, 2021. https://www.carbonbrief.org/cop26-key-outcomes-for-food-forests-land-use-and-nature-in-glasgow.

Desbureaux, Sébastien. 2021. "Subjective Modeling Choices and the Robustness of Impact Evaluations in Conservation Science." Conservation Biology35 (5): 1615–26. https://doi.org/10.1111/cobi.13728.

Fox, Jefferson, and Jean-Christophe Castella. 2013. "Expansion of Rubber (Hevea Brasiliensis) in Mainland Southeast Asia: What Are the Prospects for Small Holders?" Journal of Peasant Studies40 (January): 155–70. https://doi.org/10.1080/03066150.2012.750605.

Santika, Truly, Kerrie A. Wilson, Elizabeth A. Law, Freya A. V. St John, Kimberly M. Carlson, Holly Gibbs, Courtney L. Morgans, Marc Ancrenaz, Erik Meijaard, and Matthew J. Struebig. 2020. "Impact of Palm Oil Sustainability Certification on Village Well-Being and Poverty in Indonesia." Nature Sustainability, November, 1–11. https://doi.org/10.1038/s41893-020-00630-1.

Sellare, Jorge, Eva-Marie Meemken, Christophe Kouamé, and Matin Qaim. 2020. "Do Sustainability Standards Benefit Smallholder Farmers Also When Accounting For Cooperative Effects? Evidence from Côte d'Ivoire." American Journal of Agricultural Economics102 (2): 681–95. https://doi.org/10.1002/ajae.12015.Sellare, Jorge, Eva-Marie Meemken, and Matin Qaim. 2020. "Fairtrade, Agrochemical Input Use, and Effects on Human Health and the Environment." Ecological Economics176 (October): 106718. https://doi.org/10.1016/j.ecolecon.2020.106718.

Smith, Hilary, Juliet Lu, Phuc Xuan To, Soytavanh Mienmany, and Khonethong Soukphaxay. 2020. "Rubber Plantation Value Chains in Laos: Opportunities and Constraints in Policy, Legality and Wood Processing." Forest Trends. https://www.forest-trends.org/wp-content/uploads/2020/07/Rubber-Plantation-Value-Chains-in-Laos.pdf.

Tayleur, Catherine, Andrew Balmford, Graeme M. Buchanan, Stuart H. M. Butchart, Heather Ducharme, Rhys E. Green, Jeffrey C. Milder, et al. 2017. "Global Coverage of Agricultural Sustainability Standards, and Their Role in Conserving Biodiversity." Conservation Letters10 (5): 610–18. https://doi.org/10.1111/conl.12314.

Appendices

List of Appendix Tables

List of Appendix Figures

Appendix 1: Chapter 2

A1.1 List of supplementary materials and additional elaboration

Please note-this list describes the supplemental materials available with the online publication of the article, which can be found at: https://doi.org/10.1016/j.ecolind.2021.107490

Table S1: Indicator coding spreadsheet

a. Indicator coding. See "Supplemental Materials Indicator List" spreadsheet for a list of the indicator coding. The total number of indicators considered is 777. Wherever possible, articles included in DeFries et al. 2017 were coded consistently with that approach. Any minor divergences are noted in Column J, under "additional notes or justification." Deciding for which sustainability pillar a given indicator is relevant can be somewhat subjective. In several cases, indicators are reported upon which provide general context about the production systems and grower characteristics, rather than an "outcome" which a certification is attempting to address. Examples of this include input costs and share of household income dependent on a certain crop. Without additional context regarding yields and livelihoods, it is difficult to call an outcome for these indicators "positive" or "negative." In these cases, indicators have been given an "Other" category and excluded from directional indicator analysis, following DeFries et al. 2017. In some cases, indicators listed as "Other" in DeFries et al. have been included within a sustainability pillar in this approach. This is primarily in cases where they are linked to a specific outcome of a certification or framed as an outcome indicator in either the relevant study, or more broadly in indicator approaches for agricultural sustainability as included in COSA 2020. A few examples where this occurred included input management practices (e.g. fertilizers and agrochemicals) as well as gender indicators regarding the proportion of decisions made by the household head versus shared decision-making. Minor coding differences were noted for about 48 indicators, but general trends in the results were not affected by these differences.
b. Environmental indicator analysis. (See Supplemental Materials Indicator list, "Environmental_indicators" tab.) Indicators were coded for two characteristics: 1. Output or outcome, and 2. Data source.
c. Social indicator analysis. See Supplemental Materials Indicator list, "Social_indicators" tab.)
d. Economic indicator analysis. See Supplemental Materials Indicator list, "Economic_indicators" tab.)

Table S2: Geographical analysis

e. Country coverage analysis. See "geographical_coverage" spreadsheet, "Country_coverage" tab. Compares country coverage according to this review with information on certified extent from Tayleur et al. 2017.

Table A1.1: PRISMA review methodology materials

Table A1.1: PRISMA 2009 Checklist (Moher et al. 2009)

Note, all page numbers in the table refer to the published document (not the dissertation).

Section/topic	**#**	**Checklist item**	**Reported on page #**
TITLE			
Title	1	Identify the report as a systematic review, meta-analysis, or both.	*p.1*
ABSTRACT			
Structured summary	2	Provide a structured summary including, as applicable: background; objectives; data sources; study eligibility criteria, participants, and interventions; study appraisal and synthesis methods; results; limitations; conclusions and implications of key findings; systematic review registration number.	*p.1*
INTRODUCTION			
Rationale	3	Describe the rationale for the review in the context of what is already known.	*pp. 2, 5-6 (section 1.1)*
Objectives	4	Provide an explicit statement of questions being addressed with reference to participants, interventions, comparisons, outcomes, and study design (PICOS).	*pp. 2, 5-6 (section 1.1)*
METHODS			
Protocol and registration	5	Indicate if a review protocol exists, if and where it can be accessed (e.g., Web address), and, if available, provide registration information including registration number.	
Eligibility criteria	6	Specify study characteristics (e.g., PICOS, length of follow-up) and report characteristics (e.g., years considered, language, publication status) used as criteria for eligibility, giving rationale.	*p.6-7 (Section 2)*
Information sources	7	Describe all information sources (e.g., databases with dates of coverage, contact with study authors to identify additional studies) in the search and date last searched.	*p.6-7 (Section 2)*
Search	8	Present full electronic search strategy for at least one database, including any limits used, such that it could be repeated.	*p.6 (Section 2.1)*
Study selection	9	State the process for selecting studies (i.e., screening, eligibility, included in systematic review, and, if applicable, included in the meta-analysis).	*pp.6-7*
Data collection process	10	Describe method of data extraction from reports (e.g., piloted forms, independently, in duplicate) and any processes for obtaining and confirming data from investigators.	*p.7 (Section 2.3)*

Data items	11	List and define all variables for which data were sought (e.g., PICOS, funding sources) and any assumptions and simplifications made.	*p.7(Section 2)*
Risk of bias in individual studies	12	Describe methods used for assessing risk of bias of individual studies (including specification of whether this was done at the study or outcome level), and how this information is to be used in any data synthesis.	
Summary measures	13	State the principal summary measures (e.g., risk ratio, difference in means).	*p.7(Section 2.3)*
Synthesis of results	14	Describe the methods of handling data and combining results of studies, if done, including measures of consistency (e.g., I^2) for each meta-analysis.	
Section/topic	**#**	**Checklist item**	**Reported on page #**
Risk of bias across studies	15	Specify any assessment of risk of bias that may affect the cumulative evidence (e.g., publication bias, selective reporting within studies).	*pp.24(Section 4.3)*
Additional analyses	16	Describe methods of additional analyses (e.g., sensitivity or subgroup analyses, meta-regression), if done, indicating which were pre-specified.	
RESULTS			
Study selection	17	Give numbers of studies screened, assessed for eligibility, and included in the review, with reasons for exclusions at each stage, ideally with a flow diagram.	*p.6 (Section 2.1) and Figure 2 (p.12)*
Study characteristics	18	For each study, present characteristics for which data were extracted (e.g., study size, PICOS, follow-up period) and provide the citations.	*Table S1*
Risk of bias within studies	19	Present data on risk of bias of each study and, if available, any outcome level assessment (see item 12).	
Results of individual studies	20	For all outcomes considered (benefits or harms), present, for each study: (a) simple summary data for each intervention group (b) effect estimates and confidence intervals, ideally with a forest plot.	*Table S1*
Synthesis of results	21	Present results of each meta-analysis done, including confidence intervals and measures of consistency.	
Risk of bias across studies	22	Present results of any assessment of risk of bias across studies (see Item 15).	
Additional analysis	23	Give results of additional analyses, if done (e.g., sensitivity or subgroup analyses, meta-regression [see Item 16]).	*Table S1, Figures 5 and 6 (pp. 20-21)*
DISCUSSION			

Summary of evidence	24	Summarize the main findings including the strength of evidence for each main outcome; consider their relevance to key groups (e.g., healthcare providers, users, and policy makers).	*Section 3.3, Section 4*
Limitations	25	Discuss limitations at study and outcome level (e.g., risk of bias), and at review-level (e.g., incomplete retrieval of identified research, reporting bias).	*p.22(Section 3.5)*
Conclusions	26	Provide a general interpretation of the results in the context of other evidence, and implications for future research.	*Section 4*
FUNDING			
Funding	27	Describe sources of funding for the systematic review and other support (e.g., supply of data); role of funders for the systematic review.	*n/a*

Appendix 2: Chapter 3

A2.1 Causal diagrams and research design

The identification strategy (and covariate selection) is based on subject matter theory translated into causal graphs, also called directed acyclic graphs (DAGs). The DAGs enable a rigorous identification analysis and assessment of any remaining bias, discussed in detail below. To construct the DAGs, we used a combination of (1) SDP project documents (peer-reviewed studies, grant and NGO reports) and (2) previous literature regarding drivers of forest and poverty change. Graph design was iterative. Initial attempts to reflect all possible variables and unidirectional relationships resulted in graphs which were too computationally intensive, i.e., we could not determine the complete list of causal and non-causal paths using standard computing, which is necessary to determine the optimal adjustment set. Thus, we simplified the graphs to the extent possible while preserving the key variable relationships. The final diagrams are shown in Figures A2.1 and A2.2 and discussed further below. Time is displayed on the graphs from left to right, with pre-test variables appearing on the left.

A2.1.1 SDPs and Poverty: Graph design

Our goal was to elucidate all common causes of all pairs of variables in the graphs, as well as the factors affecting SDP participation and sustainability outcomes uncovered in the literature review (Rohrer 2018). Causal diagrams are non-linear and non-parametric, and arrows represent causal associations in the direction of the arrow, rather than positive or negative relationships or effect sizes (Elwert 2013). Below, we outline the poverty DAG structure. We assume that SDP participation is influenced by poverty conditions in the baseline year of 2005, since many SDPs set explicit poverty reduction goals (Asian Development Bank 2004; 2007). This relationship has been found in other evaluations of community-based sustainable development interventions (Ferraro and Simorangkir 2020). We hypothesize that baseline poverty conditions were influenced by livelihoods composition (e.g. proportion of on-farm vs. off-farm livelihoods), ethnic minority composition, biophysical conditions, market accessibility, government policy, education, land tenure security, and land use change (see S. M. Martin and Lorenzen 2016; Bouahom, Douangsavanh, and Rigg 2004; Wade et al. 2015).

Several variables affecting baseline poverty also affect SDP participation. Two Asian Development Bank (ADB) proposals mention that SDP villages were selected in consultation with government authorities based on development priorities (Asian Development Bank 2004; 2008). Further, we posit that demographic factors like ethnic minority group population influenced SDP participation. Some SDPs target minority groups specifically, and there are known connections between ethnic difference and poverty in Laos (Bader et al. 2016; Asian Development Bank 2008). Finally, we anticipate that accessibility to roads, infrastructure or markets may have positively influenced SDP participation (Asian Development Bank 2007).

We also identify several factors influencing changes in poverty during the study period. This includes other poverty interventions not listed in the SDP database (either government-driven or voluntary), the village resettlement program in Laos, large-scale land use changes driven by concessions or infrastructure development, and agricultural productivity (Baird and Shoemaker 2007; Manorom, Baird, and Shoemaker 2017; Nanhthavong et al. 2020).

In earlier drafts of the graphs we encountered computing challenges when calculating the adjustment set, since the graph had millions of paths. Thus, we prioritized including arrows to represent the most plausible direct causal relationships. In cases where bi-directional associations could exist, we did not include arrows. For example, land tenure security could affect livelihoods composition if insecure tenure limits the livelihoods activities that communities choose to participate in. However, livelihoods could also alter land tenure security, as has been evidenced previously in Laos (Smith et al. 2020).

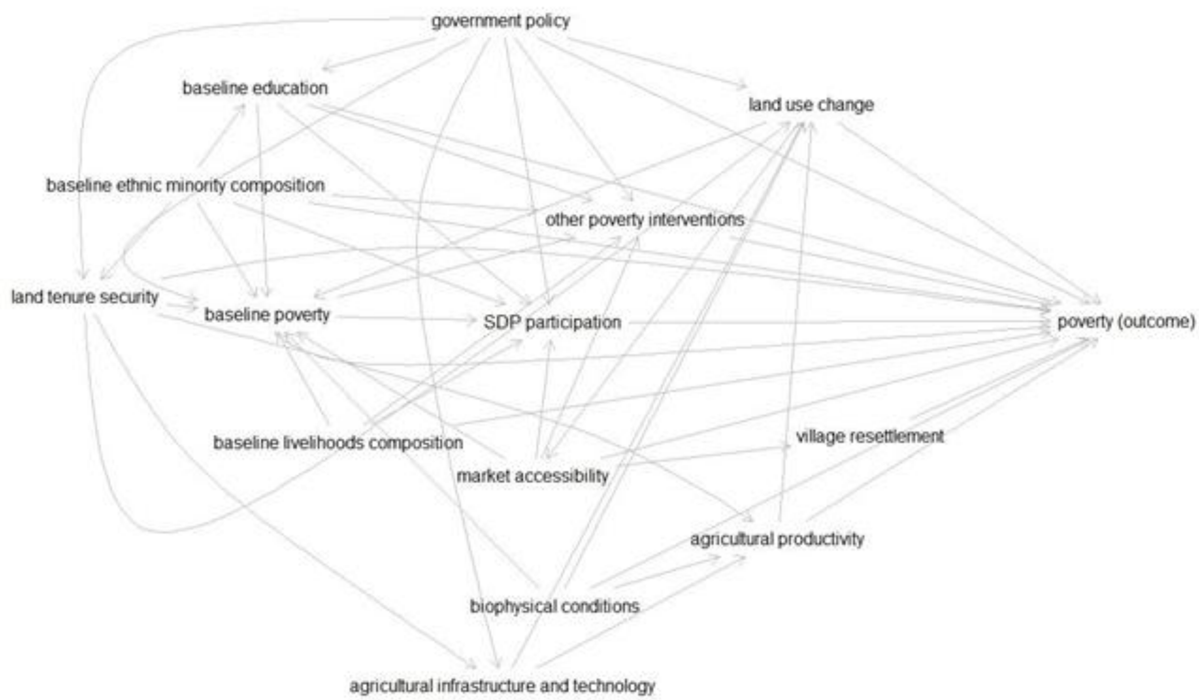


Figure A2.1: Causal diagram depicting the key causal path between SDP participation and poverty, and other relevant variables. Unobserved variables include government policy, land tenure security, other poverty interventions, and village resettlement. Source: Authors' elaboration, using dagitty package (Textor et al. 2016).

A2.1.2 SDPs and Poverty: Determining the adjustment set

The DAGs allow us to determine the appropriate variable or ‘adjustment’ set to remove most (if not all) the confounding bias through the adjustment criterion (Textor 2015). However, since not all covariates are observed, some open paths still transmit association. DAGs use terminology of ‘closed’ and ‘open’ paths based on Pearl’s backdoor criterion (Pearl 1993). Essentially, a ‘closed’ path transmits no causal association, and an ‘open path’ transmits causal association (Pearl 2009; Elwert 2013). We utilized the final graph in Figure *S1* to determine the number of open paths given our set of observed covariates (R Core Team 2020, see R code available in the Supplemental Materials). This revealed three remaining open confounding paths and three open collider paths. Table A2.1 provides an overview. The three open confounding paths necessitate close scrutiny. They are:

1) SDP < - GP -> LT -> P

● Government policy (GP) may influence selection into SDPs (SDP); it also affects the poverty outcome via land tenure security.

2) SDP <- GP -> OPI -> P

● Government policy may also affect poverty via other poverty interventions not represented in the SDP list.

3) SDP <- GP -> P

● Government policy may have a direct confounding relationship wherein it influences which villages participate in SDPs, and also influences poverty outcomes e.g. via poverty reduction programs.

Table A2.1: Causal, confounding, and collider paths in final poverty directed acyclic graph (DAG).

	Open paths	**Closed paths**
All	7	82,611
Causal	1	0
Confounding	3	667
Collider	3	81,944
Source: Authors' elaboration.		

The mechanisms by which government policy influences SDP selection and poverty outcomes may not be identical across villages. For example, consider a scenario in which the government prioritizes a certain village for poverty reduction interventions and thus also facilitates its inclusion in an SDP. This could lead to an overestimation of the treatment effect if the unobserved government policy independently contributes to poverty reduction. The same logic applies to the unobserved confounder of other poverty interventions. On the other hand, perhaps unobserved government policy worsens a village's poverty outcome (for example, a shifting cultivation reduction policy if alternative livelihood options are not secured), while the SDP improves it – this would lead to an underestimation of the treatment effect.

We argue that this confounding bias is minimized for the following reasons: first, observed covariates including baseline poverty, education, and ethnic minority composition are likely suitable proxies for government policies and other poverty interventions, for example those related to poverty reduction or agricultural intensification. Second, both government policy and land tenure security sit on open collider paths. Third, we have a strong diversity of observed measures in our poverty outcome model, including variables related to biophysical conditions, agricultural productivity, socioeconomic status, and land use change, among others. Finally, we assess the importance of any remaining source of bias through sensitivity checks, described further in *Section A2.6*, and demonstrate that results are robust to minor sources of bias.

A2.2.1 SDPs and Forest Cover: Graph Design

We assume that SDP participation is related to forest cover conditions in the baseline year of 2005, with more heavily forested areas more likely to participate. Baseline forest cover is influenced by historical forest protection measures (e.g. protected areas), land tenure security, and characteristics like market accessibility (e.g. elevation and road proximity), long-term agricultural productivity and management practices (Phompila et al. 2017; Ashraf, Pandey, and de Jong 2017; Stanley et al. 2013; World Wide Fund For Nature 2013; Wade et al. 2015; Global Environment Facility 2004; Asian Development Bank 2004). Biodiversity richness was also mentioned as influencing selection for some forest protection SDPs, as was historical donor support or NGO engagement (Global Environment Facility 2004).

Similarly to our poverty graph, socioeconomic status influences selection into SDPs. In addition to SDP participation, important influencers of forest cover will range during the study period include government land use planning (e.g. the Land and Forest Allocation program, concession moratoria), plantation and infrastructure development, land tenure security, and other resource extraction activities including illegal logging (Vongvisouk et al. 2016; Junquera et al. 2020; Thomas 2015).

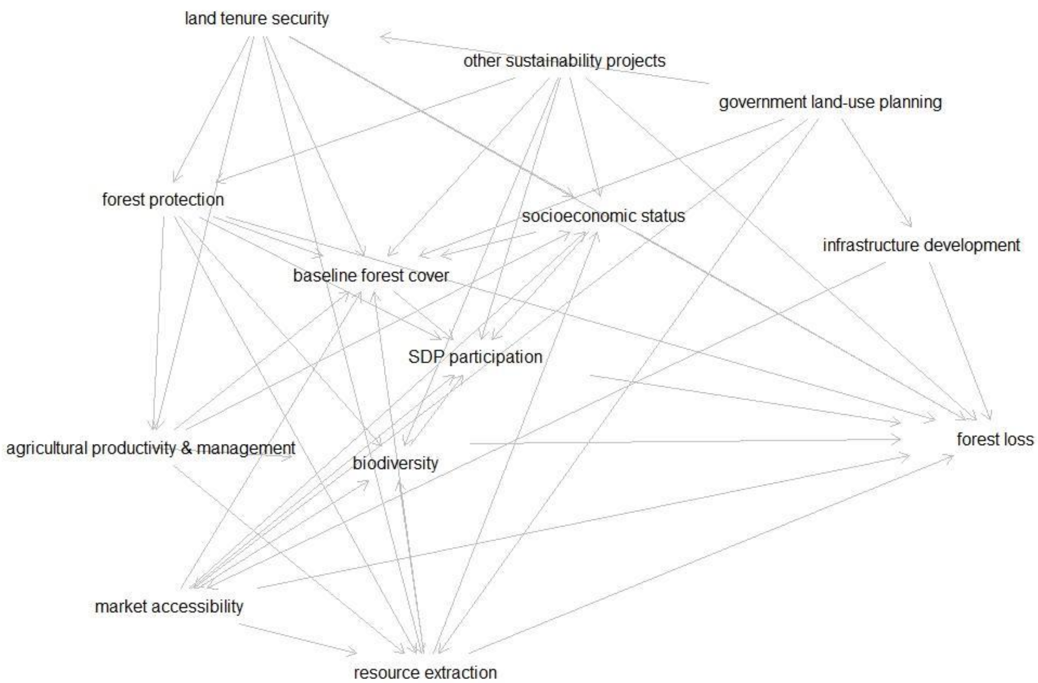


Figure A2.2: Causal diagram depicting the key causal path between SDP participation and forest cover, and other relevant variables. Government land use planning, land tenure security, resource extraction, and other sustainability efforts are unobserved. Source: Authors' elaboration, using dagitty package.

A2.2.2 SDPs and Forest Cover: Determining the adjustment set

As with the poverty graph, we utilized the final graph in Figure A2.2 to determine the number of open paths given our set of observed covariates. This revealed one remaining open confounding path and 40 open collider paths. Table A2.2 provides an overview. The remaining open path is:

(1) SDP <- OS -> FL

- Other sustainability projects may have a direct confounding relationship wherein they influence which villages participate in SDPs, and also influence forest cover outcomes.

Table A2.2: Causal, confounding, and collider paths in final forest cover directed acyclic graph (DAG).

	Open paths	**Closed paths**
All	42	165,020
Causal	1	0
Confounding	1	774
Collider	40	164,276
Source: Authors' elaboration.		

The unobserved confounder of other sustainability projects may lead to an overestimation of the treatment effect. Villages with existing sustainability projects may have been more likely to participate in SDPs, and other sustainability projects may positively influence forest cover.

We argue that this confounding bias is minimized for the following reasons: first, observed covariates including baseline forest cover, biodiversity, and protected areas are likely significantly associated to the presence of other sustainability projects based on the literature review. Second, the 'other sustainability projects' variable sits on 40 open collider paths. Third, we have a strong diversity of observed measures in our forest cover outcome model, including variables related to biophysical conditions, agricultural productivity, socioeconomic status, and land use change, among others. As mentioned above, we assess the importance of these remaining sources of bias through sensitivity checks, described further in *A2.6*.

A2.2 Data sources – additional information

A2.2.1 Outcome variables

Table A2.3 provides an overview of the outcome variables used in the analysis and the processing approach.

Table A2.3: Overview of outcome variables.

Variable	**Data Source**	**Processing notes**
2015 log of % poverty headcount	Laos Population and Housing Census (2015)	Calculated log due to right-skewed distribution
Proportional forest cover	Mapped using the approach of Potapov et al. 2019; Turubanova et al. 2018	Based on Landsat ARD; 2015 forest cover divided by land area. Forest is defined as TCC>=20% and TCH>=5m.

A2.2.2 Covariates

All matching variables were aggregated to 2015 village land-use area polygons (see A2.2.3). This typically involved calculating the weighted mean for 2015 village polygons, although in some cases the sum was calculated. Table A2.4 outlines the processing approach for estimating observed covariates.

Table A2.4: Processing approach for calculating observed covariates

Causal diagram variable	Analysis variable	Data source	Processing notes
Baseline livelihoods composition	Percentage of households with farmland in 2005	Population and Housing Census (2005, via Lao DECIDE Info Project 2020)	Geospatial linking described in A2.2.3, plus weighted mean if needed (‘05 values calculated for 2015 villages using weighted mean if there are multiple ‘05 villages in 2015 polygon)
Baseline livelihoods composition	Percentage of households with main activity non-farm sector in 2005		
Baseline poverty	Baseline poverty incidence (%)		
Baseline ethnic minority composition	% population of Lao ethno-linguistic group		
Baseline ethnic minority composition	% population of Khmuic ethno-linguistic group		
Baseline ethnic minority composition	% population of Hmong ethno-linguistic group		
Baseline ethnic minority composition	% population of Thai ethno-linguistic group		
Baseline education	% primary school completion		

Market accessibility	Baseline population density		
Market accessibility	Mean travel time to province capital		
Market accessibility	Proximity (of village centroid) to primary roads	World Food Programme/OpenStreetMap (World Food Programme and Open Street Map 2018)	Processed in QGIS using *NNjoin* plug-in. Note: no publicly available spatial data identified on roads at baseline. Addressed by visually checking WFP/OSM 2018 spatial data against a previous ADB roadmap (Asian Development Bank 2006 in Tong 2009); the majority of primary roads existed in 2006.
Land use change/infrastructure development	Number of dams in village	Mekong Region Futures Institute 2021	Processed in QGIS. Planned and proposed dams were deleted, then the number of dams per village polygon was counted. Eliminated any dams with commission dates after 2005, or unknown commission dates.
Biophysical conditions/agricultural productivity	Soil type (proportion of village area w. given soil type)	Food and Agriculture Organization of the United Nations 2007	Processed in QGIS by calculating % area of each soil type by 2015 polygon, then grouped by 2015 village ID in R.
Agricultural productivity	Long-term monthly rainfall, dry season	Wordclim (1960-2018; Harris et al. 2014)	Processed in R using packages *sf*, *raster*, *exactextractr, R.utils, and rgdal.* Created a raster stack using the Worldclim precipitation data, then calculated the area weighted mean.
Agricultural productivity	Long-term monthly rainfall, wet season	Wordclim (1960-2018; Harris et al. 2014)	
Agricultural infrastructure and technology	Proportion of land equipped for irrigation	FAO Global Map of Irrigated Areas (Siebert et al. 2013)	Processed in QGIS using the sum of area (ha) equipped for irrigation divided by the total area of the village polygon. Although published in 2013, this data reflects the baseline year of 2005 (approximately).
Biophysical conditions	Nutrient availability (mean nutrient availability status)	Harmonized World Soil Database (Fischer et al. 2008)	Processed in QGIS using zonal statistics to calculate average values per village polygon. Although the publication year is

			2008, the dataset reflects a compilation of many individual databases and it also does not have an explicit temporal component. Other sources have suggested a nominal year of 2000 based on input data (Wieder 2014).
Biophysical conditions/market access	Average elevation	SRTM (2002)	Processed in R using the packages *raster*, *exactextractr*, and *sf*, based on SRTM 90 m resolution data. Weighted mean elevation was calculated for 2015 village polygons.
Biophysical conditions/market access	Average slope	SRTM (2002)	
Baseline forest cover	Proportion forest cover	Calculated using LANDSAT Analysis-Ready Data, following Potapov et al. 2019; Turubanova et al. 2018	Calculated by dividing annual forest area by land area in 2005. Forest is defined as TCC>=20% and TCH>=5m.
Forest protection	Proportion of village located in a protected area at baseline	ICEM 2003; Knowledge for Development 2022; Open Development Mekong 2010; UNEP-WCMC 2022	Calculated in QGIS by intersecting the village and protected area polygons, calculating the area of the intersected features, then dividing by total area. There were discrepancies in spatial data regarding protected areas at baseline (2005). Therefore, we cross-checked five data sources: (1) Open Development Mekong 2010, (2) a list of protected areas and their establishment dates provided by the Department of Forestry at the National University of Laos, (3) UNEP-WCMC 2022, (4) The Knowledge for Development (K4D) website (https://apps.k4d.la/explorer/), and (5) ICEM 2003 to finalize a list of 22 protected areas existing at baseline.
Forest protection	Distance of village to protected area at baseline	ICEM 2003; Knowledge for Development 2022; Open	Processed in QGIS using *NNjoin* plugin, comparing village centroid points to the

		Development Mekong 2010; UNEP-WCMC 2022	nearest protected area.
Biodiversity	Mean village biodiversity value in 2005	Biodiversity Intactness Index (BII) (Newbold et al. 2016)	Processed in R using the packages *raster and sf*, "fast_extract" function.

A2.2.3 Village-level land-use areas and geospatial linking procedure

There are no nationally comprehensive official village boundaries available for Laos. Thus, we utilize 2015 village-level land-use area polygons, a division of the total land area into individual village's likely land use space following the method of Epprecht et al. 2008, Messerli et al. 2008, and the 2015 Population and Housing Census. Polygons were calculated using GPS coordinates of village centers from the Population and Housing Census, considering factors like slope and road quality, and using an equal travel time framework (Epprecht et al. 2008). These polygons reflect "land-use areas" rather than official administrative boundaries.

Additionally, each village is identified by a unique ID code. However, several changes in village codes occurred during the study period due to resettlement, data management, or village consolidation and division (Epprecht et al. 2018). Therefore, a geospatial linking procedure is used to match baseline data with outcome data aggregated to the polygon level, following the method used in Traldi et al. 2022. The first step involved using the numeric IDs to link most villages. Then, we used a one-to-one spatial join in QGIS to associate any "orphan" 2005 village coordinates to their 2015 overlapping polygons. Finally, there were fewer villages included in the 2015 Housing Census than 2005, presumably due to consolidation, in addition to the exclusion of Xiasomboun in our data. For cases in which multiple 2005 village points fell within a 2015 polygon, we performed a weighted average calculation of the baseline variable, using 2005 population as the weight.

Table A2.5: List of SDPs

Project Name	Start Date	End Date	Donor Name	Implementing Agency
Improve Allocate and Management of National Forest Conservation Project	1/4/2009	4/30/2011	WWF	WWF
Modern Water-Saving Irrigation	1/4/2009	4/30/2011	WWF	WWF
Agriculture Development in Upland area	1/4/2009	4/30/2011	WWF	WWF
Food Security in YothAou District Phongsaly Province	1/4/2009	4/30/2011	WWF	WWF
Improve Food Security by Irrigation for Rice Plantation Project	1/4/2009	4/30/2011	WWF	WWF
Northern Community Managed Irrigation	1/4/2009	4/30/2011	WWF	WWF
Community-Based Project to Strengthen Household Resilience to Food Price	1/4/2009	4/30/2011	WWF	WWF
Improve Food Production Project	1/4/2009	4/30/2011	WWF	WWF
Land and Forest Conservation for Sustainable Livelihood Improvement Project	1/4/2009	4/30/2011	WWF	WWF
Food Security and Nutrition in Phongsaly Province	1/4/2009	4/30/2011	WWF	WWF
Reform of the Luangprabang Agriculture and Forestry College	9/30/2010	5/30/2011	Austria	MAF
Organic Rice Value Chains Under the Profil II Project	9/30/2010	5/30/2011	Austria	MAF
Promotion of Organic Farming and Marketing in Lao PDR	9/30/2010	5/30/2011	Austria	MAF
Rural Development by Community Project Phase 2	9/30/2010	5/30/2011	Austria	MAF
Project Multi Development at Salavan Province	9/30/2010	5/30/2011	Austria	MAF
Strengthening Rattan Industry Based on Sustainable Rattan harvesting and Production in Lao PDR	9/30/2010	5/30/2011	Austria	MAF
Alternative Livelihoods for Upland Ethnic Groups in Houaphanh Province	9/30/2010	5/30/2011	Austria	MAF
Developing Carbon and Biodiversity assets for multifunctional Landscapes in the Upper Mekong	9/30/2010	5/30/2011	Austria	MAF
Food Security for Women and Rural Poor in Vienkham District, Luangprabang Province	6/1/2008	6/1/2011	Oxfam Hongkong	Oxfam Hongkong
Upland Research and Capacity Development Programme	6/1/2008	6/1/2011	Oxfam Hongkong	Oxfam Hongkong
Enhancing on-farm incomes through improved silvicultural management of teak and paper mulberry plan	6/1/2008	6/1/2011	Oxfam Hongkong	Oxfam Hongkong

Project Name	Start Date	End Date	Donor Name	Implementing Agency
Food Security Village Strengthened Project in Thathom District of Xiengkhouang Province, Lao PDR	6/1/2008	6/1/2011	Oxfam Hongkong	Oxfam Hongkong
Forestry Protection by Conservation Environment and Livelihood in Rural Area at Huaphan Province	6/1/2008	6/1/2011	Oxfam Hongkong	Oxfam Hongkong
Food Security Project in Long District, Luangnamtha Province	6/1/2008	6/1/2011	Oxfam Hongkong	Oxfam Hongkong
Best Practice Health and Husbandry of Cattle and Buffalo in the Lao PDR	6/1/2008	6/1/2011	Oxfam Hongkong	Oxfam Hongkong
Laos Extension for Agriculture Project, Phase IV	8/1/2008	7/31/2011	European Commission	CCL
Rice Production Improvement Project	8/1/2008	7/31/2011	European Commission	CCL
Rural Development in Pha-oudom, Borkeo Province	8/1/2008	7/31/2011	European Commission	CCL
Demonstration of Goat Farming Project	8/1/2008	7/31/2011	European Commission	CCL
Pig Farming in Donevay Village, Soukhouma District, Chamasack Province	8/1/2008	7/31/2011	European Commission	CCL
Local Small-scale Chicken Production to increase income of poor families in Nongdong, Phonkham and Houana Villages, Pak Ngum District, Vientiane Capital	8/1/2008	7/31/2011	European Commission	CCL
Pig Raising Demonstration to increase income of poor families in Xiengda Village, Xaysttha District,	8/1/2008	7/31/2011	European Commission	CCL
Local Small-Scale Pig Production to increase income of poor families in Namieng, Nongboua, and Napor	8/1/2008	7/31/2011	European Commission	CCL
Improvement of Food Availability and Family income generation at household level in Houaxien village, Saythany District, Vientiane Capital	8/1/2008	7/31/2011	European Commission	CCL
Pig Raising in Mai Village, Phonthong District, Vientiane Province	8/1/2008	7/31/2011	European Commission	CCL
Pig Raising to Assist Affected households, letthong Village, Phoukout District, Xiengkhouang Provinc	8/1/2008	7/31/2011	European Commission	CCL
Pig Production to Increase Low incomes household in Moung Noi Village Xaysttha District, Vientiane Capital	8/1/2008	7/31/2011	European Commission	CCL

Project Name	Start Date	End Date	Donor Name	Implementing Agency
Sustainable Forestry and Rural Development Project	8/1/2008	7/31/2011	European Commission	CCL
The Continuation of Wetlands Alliance Project	8/1/2008	7/31/2011	European Commission	CCL
Food Security in Khammuane Province	8/1/2008	7/31/2011	European Commission	CCL
Laos Bear Conservation Program	8/1/2008	7/31/2011	European Commission	CCL
Agriculture Development and Forestry in Phontong, Luangprabang Province	8/1/2008	7/31/2011	European Commission	CCL
Participatory to Crocodile Conservation Project	8/1/2008	7/31/2011	European Commission	CCL
Development of Irrigation Schemes in Champasack	8/1/2008	7/31/2011	European Commission	CCL
Northern Upland Rice Based Farming Systems Research Project	8/1/2008	7/31/2011	European Commission	CCL
Rural Community Land project	8/1/2008	7/31/2011	European Commission	CCL
Small Agro Enterprises Development for the Uplands Phase III	8/1/2008	7/31/2011	European Commission	CCL
Food Security and Nutrition in Thathom District, Xiengkhuang Province	8/1/2008	7/31/2011	European Commission	CCL
Extension to Rice Harvest one Seed Project	8/1/2008	7/31/2011	European Commission	CCL
Elds Deer Conservation Project, Xonbuly District Savannakhet Province	8/1/2008	7/31/2011	European Commission	CCL
Food Security in Xaybuathong District Khammuane Province	8/1/2008	7/31/2011	European Commission	CCL
Food Security in Dakcheung district, Sekong Province	8/1/2008	7/31/2011	European Commission	CCL
The Project Developing Improved Farming and Marketing Systems in Rainfed regions of Southern Lao PDR	8/1/2008	7/31/2011	European Commission	CCL

Project Name	Start Date	End Date	Donor Name	Implementing Agency
Northern Region Sustainable Livelihood Through Livestock Development	8/1/2008	7/31/2011	European Commission	CCL
Support to Vulnerable Households Food security in nutrition in xienghone Hongda district	6/29/2010	9/28/2011	Region Rhone Alpes	Triangle
The Project on Developing Multi-Scale climate change adaption strategies for Framing Communities in Cambodia, Lao PDR, Bangladesh and India	6/29/2010	9/28/2011	Region Rhone Alpes	Triangle
Poverty Reduction and Food Security in Muang Sepon Project	6/29/2010	9/28/2011	Region Rhone Alpes	Triangle
Participatory Land and Forestry Management Project for Reducing Deforestation	6/29/2010	9/28/2011	Region Rhone Alpes	Triangle
Development plateau of Agriculture and livelihood of Phonexay-kang Village, PhaOudom District, Bokeo Province Project	6/29/2010	9/28/2011	Region Rhone Alpes	Triangle
The Bampoo Sector Development Project, Houaphanh Province	6/29/2010	9/28/2011	Region Rhone Alpes	Triangle
Poverty Alleviation in Remote Upland Areas Project Phase II	6/29/2010	9/28/2011	Region Rhone Alpes	Triangle
Avoidance of Deforestation and Forestry Degradation along the border in Central Vietnam and Southern	6/29/2010	9/28/2011	Region Rhone Alpes	Triangle
Sustainable Food Security Development Project	6/29/2010	9/28/2011	Region Rhone Alpes	Triangle
Integrated Ecosystem and Wildlife Management Project in Boilkhamxay Province	6/29/2010	9/28/2011	Region Rhone Alpes	Triangle
Enhancing district delivery and management of agriculture extension in Lao PDR	7/15/2004	10/31/2011	ADB	Department of Irrigation
Participatory Irrigation Development	7/15/2004	10/31/2011	ADB	Department of Irrigation
Nam Et-Phou Louey National Protected Area Management and Tiger Conservation Project	7/15/2004	10/31/2011	ADB	Department of Irrigation
Project for Security and Livelihood Improvement through Sustainable Agriculture and Community Based Natural Resource Management	7/15/2004	10/31/2011	ADB	Department of Irrigation

Project Name	Start Date	End Date	Donor Name	Implementing Agency
Improved Food Diversity Project	7/15/2004	10/31/2011	ADB	Department of Irrigation
The Agro Biodiversity Initiative in Northern Part of Lao PDR	7/15/2004	10/31/2011	ADB	Department of Irrigation
Reform of the Luangprabang Agriculture and Forestry College Phase 2	7/15/2004	10/31/2011	ADB	Department of Irrigation
Livestock and Aquaculture Development for pet 18 years Project	7/15/2004	10/31/2011	ADB	Department of Irrigation
Source: Center for Development and Environment of the University of Bern (CDE).				

A2.3 SDP Typology – additional information

Table A2.6 provides additional information on the SDP typology, including example services or activities and general theories of change. Three projects which did not align with these groupings were removed from the analysis - one focused on rodents and pathogens, and two focused on aquaculture and fisheries. Theories of change may vary slightly among SDP types, but all have the potential to impact forest cover and/or poverty, either directly or indirectly. There may also be some overlap within SDPs. Theories of change are based on the literature review and review of SDP project documents.

Table A2.6: SDP Typology

Project typology	**Projects (n = 72)**	**Focus**	**Example services or activities**	**General theory of change**
Capacity development (not modelled separately due to small sample size)	5	Improving agricultural or forestry extension and/or research	improved educational institutions or training programs for ag/forestry extension	Increasing research & extension capacity will lead to more efficient and productive ag and forestry sectors, with benefits for communities who will receive a higher quality of support, as well as benefits for ecosystems when they are more effectively managed.
Food security and livelihoods development	49	Fostering food security, poverty reduction, and/or livelihood diversification, including irrigation improvement, agro-enterprise development, and livestock	provision of various inputs/resources/infrastructure and training to implement new livelihoods strategies, e.g. new crop mixes, improved irrigation	Improving food security, productivity and livelihood diversification will enhance resilience to shocks, and lead to human well-being improvements. Indirectly, this will disincentivize shifting cultivation.
Sustainable forest management	9	Forest or forestry management, including plantations (bamboo, teak) & protected forest management	training to improve smallholder forestry systems; increasing enforcement of protected areas	Enhancing the productivity and profitability of forest plantations will increase economic gains for local communities. Improving protected area management will result in reduced deforestation.

Sustainable resource management	9	Habitat conservation, either in specific place or broader landscape	research to identify opportunities for shared progress on livelihoods and resource management objectives; provision of incentives to encourage sustainable landscape management	Conservation and restoration activities will result in less deforestation and mitigate threats to biodiversity.
Source: Author's elaboration based on publicly available information *re:* SDPs in Lao during the study period.				

A2.4 Discussion of casual assumptions

IPWR was deemed appropriate for this case study based on the method's causal assumptions, which include **unconfoundedness** (also called "ignorability") and **overlap.** Unconfoundedness necessitates that the treatment selection is independent of potential outcomes after a set of covariates is conditioned upon (Schafer and Kang 2008; Rubin 1978). The overlap assumption requires that treatment and control groups have a region of common support on the propensity score; in other words, the probability of receiving treatment is between 0 and 1 for both treatment and control groups (Imbens 2003; Steiner and Cook 2013). An additional relevant assumption is the stable unit treatment value assumption (SUTVA), which states that the outcomes of one participant are not influenced by treatment assignment of any other participants (Stuart 2010; Morgan and Winship 2014). We assess the plausibility of these causal assumptions in the case of SDPs in Laos through exploration of a few examples below.

The "CarBi REDD" SDP focused on avoiding deforestation and forest degradation in border areas of southern Laos and central Vietnam via sustainable forest management in Xe Sap National Park (World Wide Fund For Nature 2013; Stanley et al. 2013). Active in 49 villages in Xekong and Saravan, this SDP was initiated in 2011 and ended in 2015. In this case, protected areas affected both selection into the SDP as well as (presumably) sustainability outcomes related to forest cover. Thus, the proportion of the village located in a protected area and the distance of the village to protected areas are critical variables to include in the IPWR to avert threats to the unconfoundedness assumption. This is also true for other protected area management SDPs, e.g. the project focused on ecosystem and wildlife management in Bolikhamxay's Nam Kading park (Global Environment Facility 2004).

Project reports state that WWF selected ten pilot communities to engage with based on the distance to the park, the community's own interest, reliance on natural resources, and village accessibility (Stanley et al. 2013). Reliance on natural resources is not a readily available baseline variable at village level in Laos; however, we expect this to be highly correlated with baseline livelihoods composition and poverty (Simo, Kanowski, and Barney 2019; Rigg 2006). For production forest management projects, sources suggested that voluntary participation could be influenced by baseline tenure arrangements, with villages experiencing more secure tenure more likely to engage with SDPs; tenure could also influence sustainability outcomes (World Bank 2018; Robinson, Holland, and Naughton-Treves 2014). Although tenure security at baseline is unobserved, this is likely linked to demographic composition of the village, with ethnic minority groups more likely to experience insecure tenure (see e.g. Broegaard, Vongvisouk, and Mertz 2017).

The final relevant causal assumption for IPWR is the stable unit treatment value assumption (SUTVA). It is conceivable that SUTVA could be violated if, for example, “control” villages near treatment villages experienced positive spillover effects from information-sharing through local social or industry networks (Jones and Gibbon 2011; Heilmayr, Carlson, and Benedict 2020; Pfaff and Robalino 2017). This could result in an underestimation of the treatment effect. Negative spillovers for forest cover outcomes - or leakage - could also occur, for example if deforestation within SDPs is merely re-located to non-SDP villages (Pfaff and Robalino 2017); this could result in an overestimation of the treatment effect. Finally, the different SDP types - and therefore, different treatment intensities and mechanisms - necessitate the exploration of effect heterogeneity; this is elaborated upon further in the Results section in the main text.

A2.5 Propensity score and ATT estimation

A2.5.1 Propensity score estimation and doubly robust regression

As described in the main text, our analysis includes 8 models (Table A2.7). The models use the master covariate list described in Table A2.7. We estimate the propensity score via logistic regression of treatment Z on observed predictors β (1):

$$logit(Z) = \alpha + \beta$$

Then we conduct weighted least squares regression with covariates (Steiner and Cook 2013) (2):

$$Y = \alpha + \tau Z + X'\beta + \varepsilon$$

where Z is the treatment indicator, τ the ATT, X the vector of mean centered covariates of the adjustment set, and β the corresponding coefficient vector.

Table A2.7: Model Overview.

Model outcome	**Treatment group**	**# of covariates, propensity score model**	**# of additional covariates in doubly adjusted regression**
2015 poverty headcount (%)	All SDPs (n=1,452)	32	21
	SDP Type 1/Food security and livelihoods SDPs[22] (n=829)		
	SDP Type 3/ Sustainable forest management SDPs (n= 166)		
	SDP Type 4/Sustainable resource management\ (n= 375)		

[22] SDP Type 2 is capacity development, which we do not include as a stand-alone model due to small sample size.

Proportional forest cover in 2015	All SDPs (n=1,452)	32	21
	SDP Type 1/Food security and livelihoods SDPs (n=829)		
	SDP Type 3/ Sustainable forest management SDPs (n= 166)		
	SDP Type 4/ Sustainable resource management (n= 375)		

The following variables were used for the propensity score estimation and matching:

(1) proportion of the area protected, (2) distance to protected areas, (3) distance to roads, (4) percent households with farmland, (5) percent households with nonfarm livelihoods, (6) baseline poverty, (7) proximity to province capital, (8) baseline population density, (9) baseline forest cover, (10) baseline percent population ethnolinguistic category Lao, (11) average elevation, (12) average slope, (13) average dry season long-term rainfall, (14) average wet season long-term rainfall, (15) proportion of the area equipped for irrigation, (16) number of dams in the village, (17) average nutrient availability, (18-27) proportional soil type variables, (28) baseline percent population ethnolinguistic category Khmuic, (29) baseline percent population ethnolinguistic category Thai, (30) baseline percent population ethnolinguistic category Hmong, (31) mean biodiversity at baseline, and (32) percent of the population which had completed primary school.

The following variables were used as additional covariates for the doubly adjusted regression:

(1) proportion of the area protected, (2) distance to protected areas, (3) distance to roads, (4) percent households with farmland, (5) percent households with nonfarm livelihoods, (6) baseline poverty, (7) proximity to province capital, (8) baseline population density, (9) baseline forest cover, (10) baseline percent population ethnolinguistic category Lao, (11) average elevation, (12) average slope, (13) average dry season long-term rainfall, (14) average wet season long-term rainfall, (15) proportion of the area equipped for irrigation, (16) average nutrient availability, (17) baseline percent population ethnolinguistic category Khmuic, (18) baseline percent population ethnolinguistic category Thai, (19) baseline percent population ethnolinguistic category Hmong, (20) mean biodiversity at baseline, and (21) percent of the population which had completed primary school.

A2.5.2 Matching characteristics: Love plots

Below are summary love plot visualizations for all propensity score models included in the analysis, generated by the *MatchIt* and *cobalt* packages (Ho et al. 2011; Greifer 2022). Ideally, standardized mean differences should be <.1 (Rubin 2001).

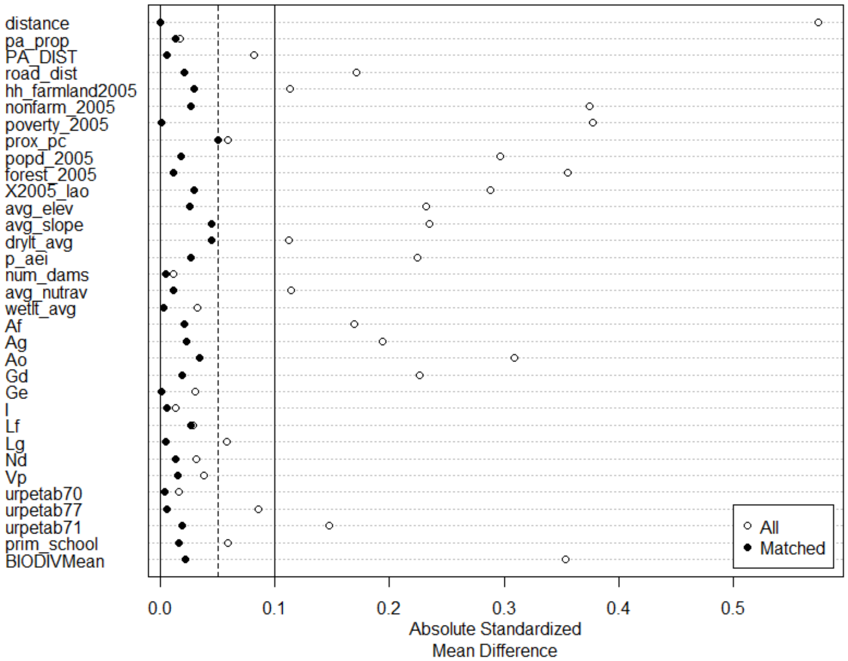


Figure A2.3: Love plot: All SDPs PS model

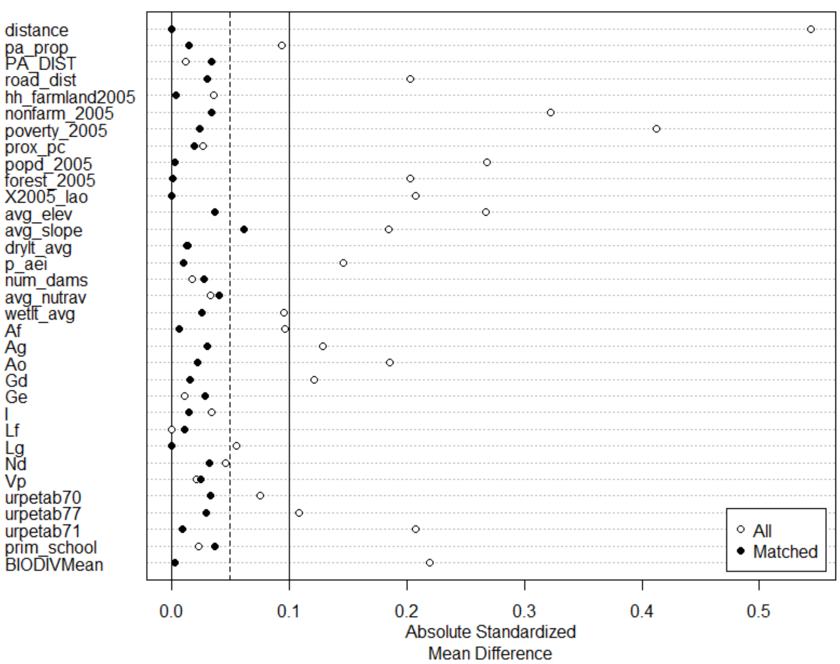


Figure A2.4: Love plot: Food security and livelihoods SDPs PS model

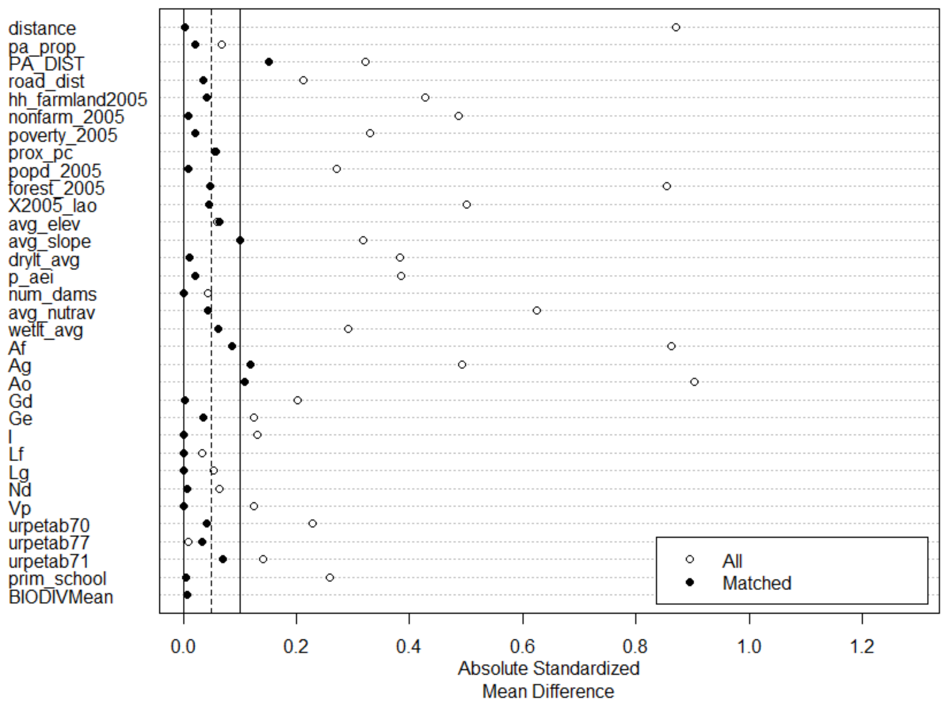


Figure A2.5: Love plot: Sustainable forest management SDPs PS model

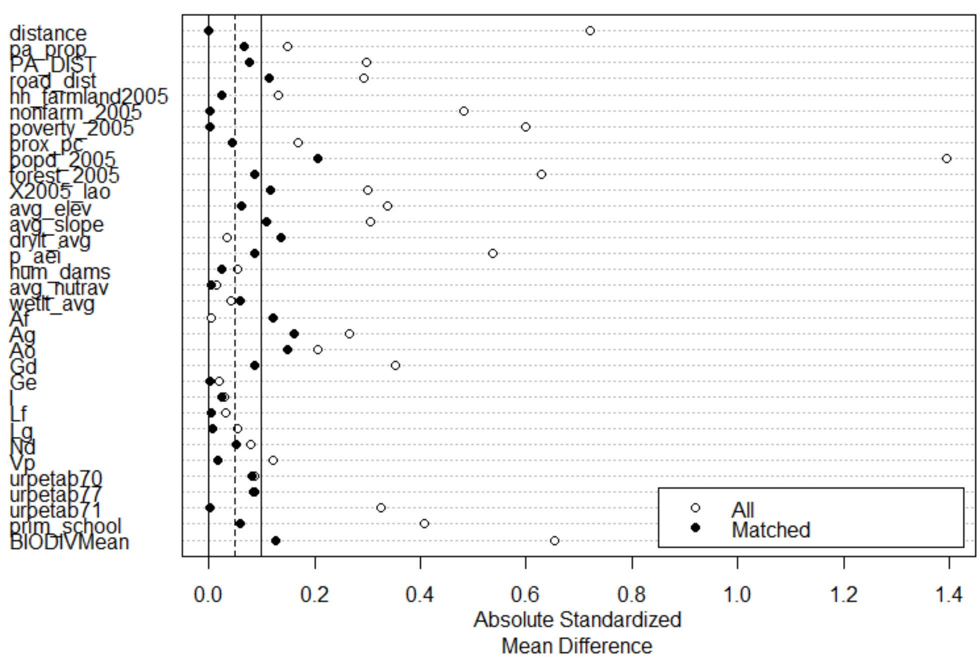


Figure A2.6: Love plot: Sustainable resource management SDPs PS model

A2.6 Sensitivity Analysis

A2.6.1 E-Values

We use function “evalues.OLS” in the *EValue* package for sensitivity analysis (Mathur et al. 2021). The value represents the minimum strength of association that an unmeasured confounder would require to fully invalidate an observed treatment effect, conditional on observed covariates (VanderWeele and Ding 2017). The larger the value, the greater the unmeasured confounding would need to be to invalidate the effect estimate (Ibid). Results of the e-values calculations are below. For the significant poverty models, results suggest that effect estimates are somewhat sensitive to unobserved confounding, with e-values ranging from 1.29 to 1.31. In other words, an unmeasured confounder associated with selection into SDPs and the poverty outcome would need to exhibit a risk ratio of >1.29 to explain away the treatment effect (VanderWeele and Ding 2017). For the significant forest cover outcome models, e-values range from 1.19 to 1.29.

Poverty estimate, All SDP Projects	**Point**	**Lower**	**Upper**
RR	0.9432962	0.9116564	0.9760341
E-Values	1.3125523	N/A	1.1831648

Poverty estimate, Food security and livelihoods projects	**Point**	**Lower**	**Upper**
RR	0.9498904	0.9099831	0.9915478
E-Values	1.2884136	N/A	1.1012437

Forest cover estimate, all SDP projects	**Point**	**Lower**	**Upper**
RR	1.025973	1.011916	1.040226
E-Values	1.189215	1.121723	N/A

Forest cover estimate, Food security and livelihoods projects	**Point**	**Lower**	**Upper**
RR	1.026832	1.010726	1.043194
E-Values	1.192819	1.114848	N/A

Forest cover estimate, Resource management projects	**Point**	**Lower**	**Upper**
RR	1.054096	1.023404	1.085709

E-Values	1.292889	1.178166	N/A

A2.6.2 Spatial Autocorrelation and Moran's I

We test for spatial autocorrelation by calculating Moran's I of the post-matching regression residuals, following Oldekop et al. 2019, and using packages *sf* and *spdep* (Pebesma et al. 2022; Bivand et al. 2022). Results are shown in Table A2.8. For the poverty models, inclusion of covariates in the regression model results in small increases in spatial autocorrelation as measured by Moran's I, but spatial autocorrelation is still relatively low for all models (Moran's I <=.04). For the forest cover models, including covariates in the regression improves the degree of spatial autocorrelation, which is relatively low for all outcome models (Moran's I <=.02).

Table A2.8: Moran's I values of model residuals

OutcomesRep	ModelType	MoranIAdjust	MoranINoAdjust	MoranIDiff
Poverty	All SDPs	0.040142362	0.021069	-0.01907
Poverty	type 1 (food security and livelihoods)	0.036439794	0.034249	-0.00219
Poverty	type 3 (forest management)	0.011013298	0.009793	-0.00122
Poverty	type 4 (resource management)	0.002727198	0.029053	0.026326
Forest cover	All SDPs	0.009495579	0.056542	0.047047
Forest cover	type 1 (food security and livelihoods)	0.023312703	0.062233	0.03892
Forest cover	type 3 (forest management)	-0.000712669	0.031066	0.031779
Forest cover	type 4 (resource management)	0.004914161	0.113457	0.108543

A2.7 Bibliography

Ashraf, Jawaid, Rajiv Pandey, and Wil de Jong. 2017. “Assessment of Bio-Physical, Social and Economic Drivers for Forest Transition in Asia-Pacific Region.” *Forest Policy and Economics*, Forest transition in Asia, 76 (March): 35–44. https://doi.org/10.1016/j.forpol.2016.07.008.

Asian Development Bank. 2004. “Northern Community-Managed Irrigation Sector Project: Report and Recommendation of the President to the Board of Directors on a Proposed Loan.” RRP: LAO 34188. Asian Development Bank. https://www.adb.org/sites/default/files/project-document/69972/rrp-lao-34188.pdf.

———. 2006. “Road Network Map of Lao PDR.”

———. 2007. “Northern Region Sustainable Livelihoods through Livestock Development Project.” 35297. Project Administration Memorandum. Asian Development Bank. https://www.adb.org/sites/default/files/project-document/66425/35297-lao-pam.pdf.

———. 2008. “Alternative Livelihood for Upland Ethnic Groups in Houaphanh Province (Financed by the Japan Fund for Poverty Reduction).” 41096. Grant Assistance Report. Asian Development Bank. https://www.adb.org/sites/default/files/project-document/65334/41096-lao-gar.pdf.

Bader, Christoph, Sabin Bieri, Urs Wiesmann, and Andreas Heinimann. 2016. “Differences Between Monetary and Multidimensional Poverty in the Lao PDR: Implications for Targeting of Poverty Reduction Policies and Interventions.” *Poverty & Public Policy* 8 (2): 171–97. https://doi.org/10.1002/pop4.140.

Baird, Ian G., and Bruce Shoemaker. 2007. “Unsettling Experiences: Internal Resettlement and International Aid Agencies in Laos.” *Development and Change* 38 (5): 865–88. https://doi.org/10.1111/j.1467-7660.2007.00437.x.

Bivand, Roger, Micah Altman, Luc Anselin, Renato Assunção, Olaf Berke, Guillaume Blanchet, Marilia Carvalho, et al. 2022. *Spdep: Spatial Dependence: Weighting Schemes, Statistics* (version 1.2-4). https://CRAN.R-project.org/package=spdep.

Bouahom, Bounthong, Linkham Douangsavanh, and Jonathan Rigg. 2004. “Building Sustainable Livelihoods in Laos: Untangling Farm from Non-Farm, Progress from Distress.” *Geoforum*, Themed section on The Spaces of Critical Geography, 35 (5): 607–19. https://doi.org/10.1016/j.geoforum.2004.02.002.

Broegaard, Rikke Brandt, Thoumthone Vongvisouk, and Ole Mertz. 2017. “Contradictory Land Use Plans and Policies in Laos: Tenure Security and the Threat of Exclusion.” *World Development* 89 (January): 170–83. https://doi.org/10.1016/j.worlddev.2016.08.008.

Elwert, Felix. 2013. “Graphical Causal Models.” In *Handbook of Causal Analysis for Social Research*, edited by Stephen L. Morgan, 245–73. Handbooks of Sociology and Social Research. Dordrecht: Springer Netherlands. https://doi.org/10.1007/978-94-007-6094-3_13.

Epprecht, Michael, Nicholas Bosoni, Albrecht Ehrensperger, Hiroko Nagasawa, Juliet Lu, David Studer, Patrizia Vollmar, Rasso Bernhard, and Vilaysook Sisoulath. 2018. *Socio-Economic Atlas of the Lao PDR. Patterns and Trends from 2005 to 2015*. Bern, Switzerland and Vientiane, Lao PDR: Centre for Development and Environment, University of Bern, and Lao Statistics Bureau, Lao PDR, with Bern Open Publishing. https://doi.org/10.7892/boris.120671.

Epprecht, Michael, Nicolas Minot, Reno Dewina, Peter Messerli, and Andreas Heinimann. 2008. “The Geography of Poverty and Inequality in the Lao PDR.” Application/pdf. https://doi.org/10.7892/BORIS.34330.

Ferraro, Paul J., and Rhita Simorangkir. 2020. “Conditional Cash Transfers to Alleviate Poverty Also Reduced Deforestation in Indonesia.” *Science Advances* 6 (24): eaaz1298. https://doi.org/10.1126/sciadv.aaz1298.

Fischer, G., F. Nachtergaele, S. Prieler, H. T. van Velthuizen, L. Verelst, and D. Wiberg. 2008. “Global Agro-Ecological Zones Assessment for Agriculture. Dataset.” Laxenburg, Austria / Rome, Italy: IIASA and FAO. https://webarchive.iiasa.ac.at/Research/LUC/External-World-soil-database/HTML/SoilQualityData.html?sb=11.

Food and Agriculture Organization of the United Nations. 2007. “FAO Digital Soil Map of the World (Version 3.6). Dataset.” 2007. https://data.apps.fao.org/map/catalog/srv/eng/catalog.search#/metadata/446ed430-8383-11db-b9b2-000d939bc5d8.

Global Environment Facility. 2004. “Integrated Ecosystem and Wildlife Management in Bolikhamxay Province - Project Proposal.” MEDIUM-SIZED PROJECT PROPOSAL. Global Environment Facility. https://documents1.worldbank.org/curated/en/395381468772454991/pdf/316500LaoPDR0MSP0Brief012113104.pdf.

Greifer, Noah. 2022. *Cobalt: Covariate Balance Tables and Plots* (version 4.3.2). https://CRAN.R-project.org/package=cobalt.

Harris, I., P.D. Jones, T.J. Osborn, and D.H. Lister. 2014. “Updated High-Resolution Grids of Monthly Climatic Observations - the CRU TS3.10 Dataset: UPDATED HIGH-RESOLUTION GRIDS OF MONTHLY CLIMATIC OBSERVATIONS.” *International Journal of Climatology* 34 (3): 623–42. https://doi.org/10.1002/joc.3711.

Heilmayr, Robert, Kimberly M. Carlson, and Jason Jon Benedict. 2020. “Deforestation Spillovers from Oil Palm Sustainability Certification.” *Environmental Research Letters* 15 (7): 075002. https://doi.org/10.1088/1748-9326/ab7f0c.

Ho, Daniel, Kosuke Imai, Gary King, and Elizabeth A. Stuart. 2011. “MatchIt: Nonparametric Preprocessing for Parametric Causal Inference.” *Journal of Statistical Software* 42 (1): 1–28. https://doi.org/10.18637/jss.v042.i08.

ICEM. 2003. “Lao People’s Democratic Republic: National Report on Protected Areas and Development.” Review of Protected Areas and Development in the Four Countries of the Lower Mekong River Region. Indooroopilly, Queensland, Australia: International Centre for Environmental Management (ICEM). https://icem.com.au/documents/biodiversity/pad/lao_pdr_nr.pdf.

Imbens, Guido W. 2003. “Nonparametric Estimation of Average Treatment Effects under Exogeneity: A Review.” t0294. National Bureau of Economic Research. https://doi.org/10.3386/t0294.

Jones, Sam, and Peter Gibbon. 2011. “Developing Agricultural Markets in Sub-Saharan Africa: Organic Cocoa in Rural Uganda.” *The Journal of Development Studies* 47 (10): 1595–1618. https://doi.org/10.1080/00220388.2011.579107.

Junquera, Victoria, Patrick Meyfroidt, Zhanli Sun, Phokham Latthachack, and Adrienne Grêt-Regamey. 2020. “From Global Drivers to Local Land-Use Change: Understanding the

Northern Laos Rubber Boom." *Environmental Science & Policy* 109 (July): 103–15. https://doi.org/10.1016/j.envsci.2020.04.013.
Knowledge for Development. 2022. "K4D Map Explorer." 2022. https://apps.k4d.la/explorer/.
Lao DECIDE Info Project. 2020. "Lao DECIDE Info - Start." Lao DECIDE Info. 2020. http://www.decide.la/en/.
Manorom, Kanokwan, Ian G. Baird, and Bruce Shoemaker. 2017. "The World Bank, Hydropower-Based Poverty Alleviation and Indigenous Peoples: On-the-Ground Realities in the Xe Bang Fai River Basin of Laos." *Forum for Development Studies* 44 (2): 275–300. https://doi.org/10.1080/08039410.2016.1273850.
Martin, Sarah M., and Kai Lorenzen. 2016. "Livelihood Diversification in Rural Laos." *World Development* 83 (July): 231–43. https://doi.org/10.1016/j.worlddev.2016.01.018.
Mathur, Maya B., Louisa H. Smith, Peng Ding, and Tyler J. VanderWeele. 2021. *EValue: Sensitivity Analyses for Unmeasured Confounding and Other Biases in Observational Studies and Meta-Analyses* (version 4.1.3). https://CRAN.R-project.org/package=EValue.
Mekong Region Futures Institute. 2021. "Dataset on the Dams of the Greater Mekong." Bangkok: Mekong Region Futures Institute.
Messerli, Peter, Andreas Heinimann, Michael Epprecht, Phonesaly Souksavath, Thiraka Chanthalanouvong, and Nicholas Minot. 2008. *Socio-Economic Atlas of the Lao PDR. An Analysis Based on the 2005 Population and Housing Census. Messerli, Peter; Heinimann, Andreas; Epprecht, Michael; Souksavath, Phonesaly; Chanthalanouvong, Thiraka; Minot, Nicholas (2008). Socio-Economic Atlas of the Lao PDR. An Analysis Based on the 2005 Population and Housing Census [Textbook] . Bern: Geographica Bernensia*. Bern: Geographica Bernensia. http://www.laoatlas.net/links/PDF%20download.html.
Morgan, Stephen L., and Christopher Winship. 2014. *Counterfactuals and Causal Inference: Methods and Principles for Social Research*. 2nd ed. Analytical Methods for Social Research. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9781107587991.
Nanhthavong, Vong, Michael Epprecht, Cornelia Hett, Julie G. Zaehringer, and Peter Messerli. 2020. "Poverty Trends in Villages Affected by Land-Based Investments in Rural Laos." *Applied Geography* 124 (November): 102298. https://doi.org/10.1016/j.apgeog.2020.102298.
Newbold, Tim, Lawrence Hudson, Andy Arnell, Sara Contu, Adriana De Palma, Simon Ferrier, Samantha L. L. Hill, et al. 2016. "Global Map of the Biodiversity Intactness Index, from Newbold et al. (2016) Science," July. https://doi.org/10.5519/0009936.
Oldekop, Johan A., Katharine R. E. Sims, Birendra K. Karna, Mark J. Whittingham, and Arun Agrawal. 2019. "Reductions in Deforestation and Poverty from Decentralized Forest Management in Nepal." *Nature Sustainability* 2 (5): 421–28. https://doi.org/10.1038/s41893-019-0277-3.
Open Development Mekong. 2010. "Laos Protected Areas and Heritage Sites- OD Mekong Datahub." 2010. https://data.opendevelopmentmekong.net/dataset/laos-protected-areas-and-heritage-sites/resource/c4333eba-f031-45cd-bd58-167709847369?inner_span=True.
Pearl, Judea. 1993. "[Bayesian Analysis in Expert Systems]: Comment: Graphical Models, Causality and Intervention." *Statistical Science* 8 (3): 266–69.
———. 2009. *Causality*. 2nd ed. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511803161.

Pebesma, Edzer, Roger Bivand, Etienne Racine, Michael Sumner, Ian Cook, Tim Keitt, Robin Lovelace, et al. 2022. *Sf: Simple Features for R* (version 1.0-7). https://CRAN.R-project.org/package=sf.

Pfaff, Alexander, and Juan Robalino. 2017. "Spillovers from Conservation Programs." *Annual Review of Resource Economics* 9 (1): 299–315.

Phompila, Chittana, Megan Lewis, Bertram Ostendorf, and Kenneth Clarke. 2017. "Forest Cover Changes in Lao Tropical Forests: Physical and Socio-Economic Factors Are the Most Important Drivers." *Land* 6 (2): 23. https://doi.org/10.3390/land6020023.

Potapov, Peter, A. Tyukavina, S. Turubanova, Y. Talero, A. Hernandez-Serna, M. C. Hansen, D. Saah, et al. 2019. "Annual Continuous Fields of Woody Vegetation Structure in the Lower Mekong Region from 2000-2017 Landsat Time-Series." *Remote Sensing of Environment* 232 (October): 111278. https://doi.org/10.1016/j.rse.2019.111278.

R Core Team. 2020. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. https://www.R-project.org.

Rigg, J. D. 2006. "Forests, Marketization, Livelihoods and the Poor in the Lao PDR." *Land Degradation & Development* 17 (2): 123–33. https://doi.org/10.1002/ldr.719.

Robinson, Brian E., Margaret B. Holland, and Lisa Naughton-Treves. 2014. "Does Secure Land Tenure Save Forests? A Meta-Analysis of the Relationship between Land Tenure and Tropical Deforestation." *Global Environmental Change* 29 (November): 281–93. https://doi.org/10.1016/j.gloenvcha.2013.05.012.

Rohrer, Julia M. 2018. "Thinking Clearly About Correlations and Causation: Graphical Causal Models for Observational Data." *Advances in Methods and Practices in Psychological Science* 1 (1): 27–42. https://doi.org/10.1177/2515245917745629.

Rubin, Donald B. 1978. "Bayesian Inference for Causal Effects: The Role of Randomization." *The Annals of Statistics* 6 (1): 34–58. https://doi.org/10.1214/aos/1176344064.

———. 2001. "Using Propensity Scores to Help Design Observational Studies: Application to the Tobacco Litigation." *Health Services and Outcomes Research Methodology* 2 (3): 169–88. https://doi.org/10.1023/A:1020363010465.

Schafer, Joseph L., and Joseph Kang. 2008. "Average Causal Effects from Nonrandomized Studies: A Practical Guide and Simulated Example." *Psychological Methods* 13 (4): 279–313. https://doi.org/10.1037/a0014268.

Siebert, Stefan, Verena Henrich, Karen Frenken, and Jacob Burke. 2013. "Global Map of Irrigation Areas Version 5. Dataset." Bonn, Germany / Rome, Italy: Rheinische Friedrich-Wilhelms-University / Food and Agriculture Organization of the United Nations. https://www.fao.org/aquastat/en/geospatial-information/global-maps-irrigated-areas/latest-version.

Simo, Alex Van Der Meer, Peter Kanowski, and Keith Barney. 2019. "Revealing Environmental Income in Rural Livelihoods: Evidence from Four Villages in Lao PDR." *Forests, Trees and Livelihoods* 28 (1): 16–33. https://doi.org/10.1080/14728028.2018.1552540.

Smith, Hilary, Juliet Lu, Phuc Xuan To, Soytavanh Mienmany, and Khonethong Soukphaxay. 2020. "Rubber Plantation Value Chains in Laos: Opportunities and Constraints in Policy, Legality and Wood Processing." Forest Trends. https://www.forest-trends.org/wp-content/uploads/2020/07/Rubber-Plantation-Value-Chains-in-Laos.pdf.

Stanley, Scott Alexander, Richard Hackman, Jeremy Ferrand, and Forest Carbon Consultants Inc. 2013. "REDD Feasibility Study for Xe Sap, Saola, and Saola Quang Nam NPAs in Lao PDR and Vietnam." World Wide Fund for Nature, Car-Bi Project.

Steiner, Peter M., and David Cook. 2013. “Matching and Propensity Scores.” In *The Oxford Handbook of Quantitative Methods*. Oxford University Press.
Stuart, Elizabeth A. 2010. “Matching Methods for Causal Inference: A Review and a Look Forward.” *Statistical Science: A Review Journal of the Institute of Mathematical Statistics* 25 (1): 1–21. https://doi.org/10.1214/09-STS313.
Textor, Johannes. 2015. “Drawing and Analyzing Causal DAGs with DAGitty.” http://www.dagitty.net/manual-2.x.pdf.
Textor, Johannes, Benito van der Zander, Mark S. Gilthorpe, Maciej Liskiewicz, and George Th Ellison. 2016. “Robust Causal Inference Using Directed Acyclic Graphs: The R Package ‘Dagitty.’” *International Journal of Epidemiology* 45 (6): 1887–94. https://doi.org/10.1093/ije/dyw341.
Thomas, Ian Lloyd. 2015. “USAID Lowering Emissions in Asia’s Forests (USAID LEAF): Drivers of Forest Change in the Greater Mekong Subregion. Laos Country Report.” USAID LEAF. https://www.uncclearn.org/wp-content/uploads/library/fao13102015_4.pdf.
Tong, Pei Sin. 2009. “LAO PEOPLE’S DEMOCRATIC REPUBLIC FORESTRY OUTLOOK STUDY.” Working Paper No. APFSOS II/WP/2009/17. Working Paper Series. Bangkok: Food And Agriculture Organization Of The United Nations.
Traldi, Rebecca, Julie A. Silva, Peter Potapov, Alexandra Tyukavina, Michael Epprecht, Meredith L. Gore, and Chittana Phompila. 2022. “Cultivating Inequality? Regional Rubber Dynamics and Implications for Voluntary Sustainability Programs in Lao PDR.” *Under Review*.
Turubanova, Svetlana, Peter V. Potapov, Alexandra Tyukavina, and Matthew C. Hansen. 2018. “Ongoing Primary Forest Loss in Brazil, Democratic Republic of the Congo, and Indonesia.” *Environmental Research Letters* 13 (7): 074028. https://doi.org/10.1088/1748-9326/aacd1c.
UNEP-WCMC. 2022. “Protected Planet.” Protected Planet. 2022. https://www.protectedplanet.net/en.
VanderWeele, Tyler J., and Peng Ding. 2017. “Sensitivity Analysis in Observational Research: Introducing the E-Value.” *Annals of Internal Medicine* 167 (4): 268. https://doi.org/10.7326/M16-2607.
Vongvisouk, Thoumthone, Rikke Brandt Broegaard, Ole Mertz, and Sithong Thongmanivong. 2016. “Rush for Cash Crops and Forest Protection: Neither Land Sparing nor Land Sharing.” *Land Use Policy* 55 (September): 182–92. https://doi.org/10.1016/j.landusepol.2016.04.001.
Wade, Len, Tamara Jackson, Jenny Locker, and Sarah Boothey. 2015. “Developing Improved Farming and Marketing Systems in Rainfed Regions of Southern Lao PDR - Final Report.” CSE/2009/004. Australian Centre for International Agricultural Research. https://aciar.gov.au/publication/technical-publications/developing-improved-farming-and-marketing-systems-rainfed-regions-southern-lao-pdr.
Wieder, W. 2014. “Regridded Harmonized World Soil Database v1.2,” 59.234908 MB. https://doi.org/10.3334/ORNLDAAC/1247.
World Bank. 2018. “Lao People’s Democratic Republic- Sustainable Forestry for Rural Development Project.” 125929. Independent Evaluation Group, Project Performance Assessment Report. Washington DC: The World Bank.

https://documents1.worldbank.org/curated/fr/239161528225963013/pdf/125929-PPAR-P163409-P064886-PUBLIC.pdf.

World Food Programme, and Open Street Map. 2018. "Laos Road Network (Main Roads). Dataset." WFP GeoNode. 2018. https://geonode.wfp.org/layers/ogcserver.gis.wfp.org%3Ageonode%3Alao_trs_roads_osm.

World Wide Fund For Nature. 2013. "Results from the CarBi REDD Feasibility Study | WWF." World Wide Fund For Nature. 2013. https://wwf.panda.org/wwf_news/?208939/Results-from-the-CarBi-REDD-Feasibility-Study.

Appendix 3: Chapter 4

A3.1 Overall consumption per capita, rural areas (LECS 4, 2007/08-LECS 5 2012/13)

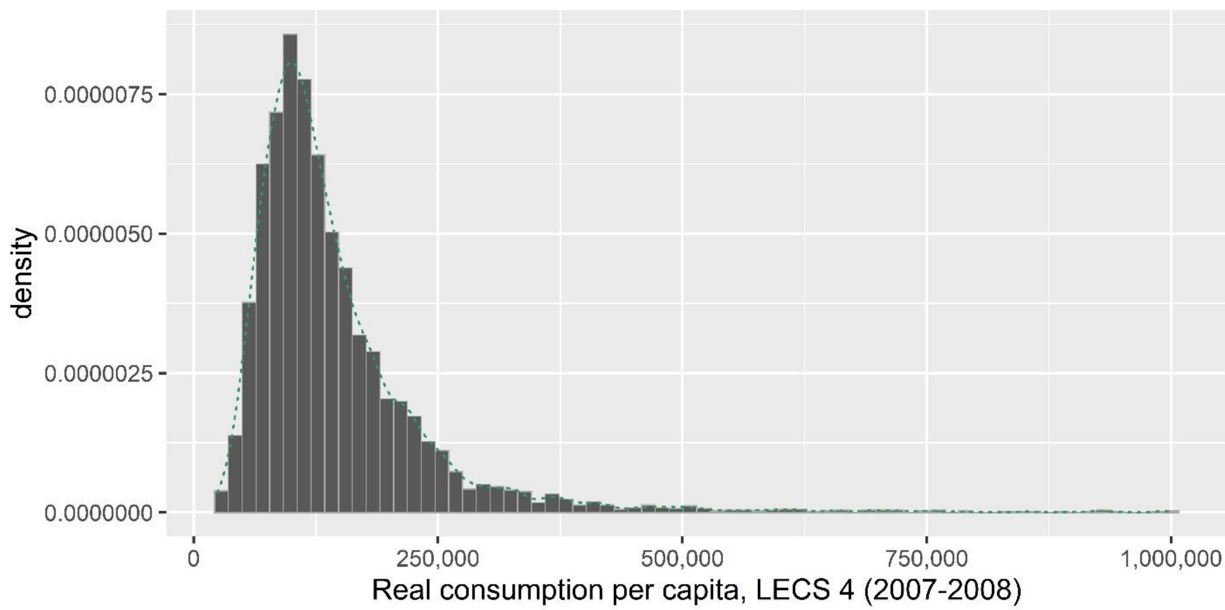


Figure A3.1: Real consumption per capita (kip), LECS 4 (Histogram with density curve; outliers >1,000,000 kip removed for visualization purposes). Source: Authors' estimates using the LECS4 and LECS5 Household Survey Data.

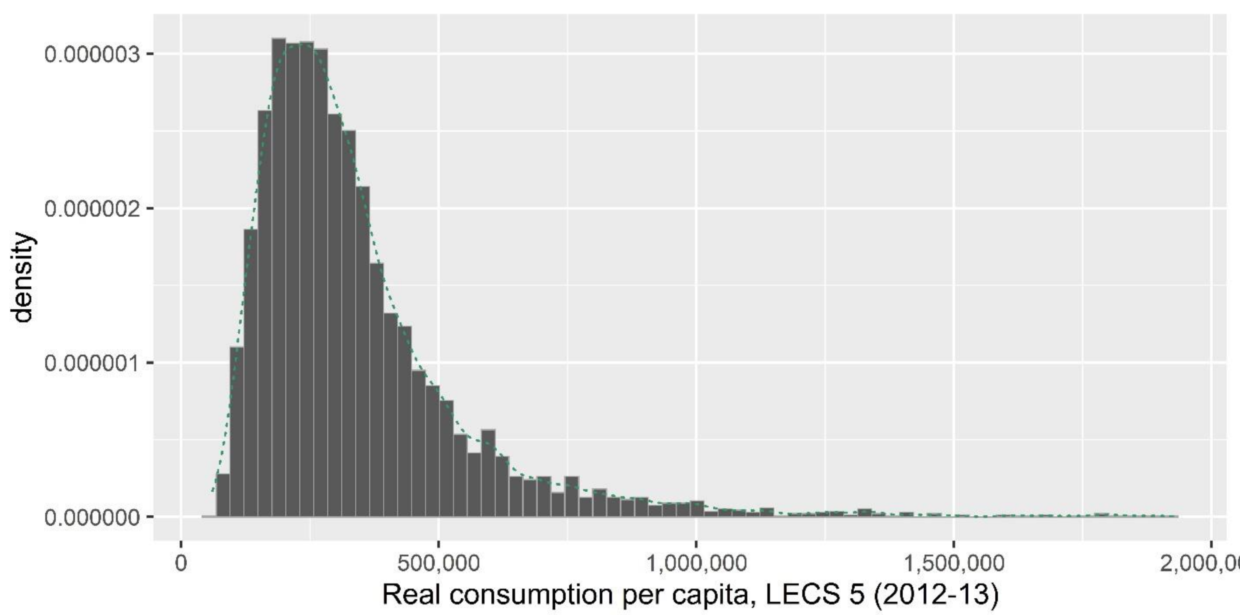


Figure A3.2: Real consumption per capita (kip), LECS 5 (Histogram with density curve; outliers of >2,000,000 kip removed for visualization purposes). Source: Authors' estimates using the LECS4 and LECS5 Household Survey Data.

A3.2 Additional elaboration-ethnic minority groups

This section provides some additional elaboration on the expenditure distribution for ethnic minority groups between 2007/08-2012/13, including real consumption per capita (Figure A3.3), Gini decomposition results (Table A3.1), and Z-K index results (Table A3.2). While not the primary focus of this paper, these results do provide some interesting context, and thus are included here.

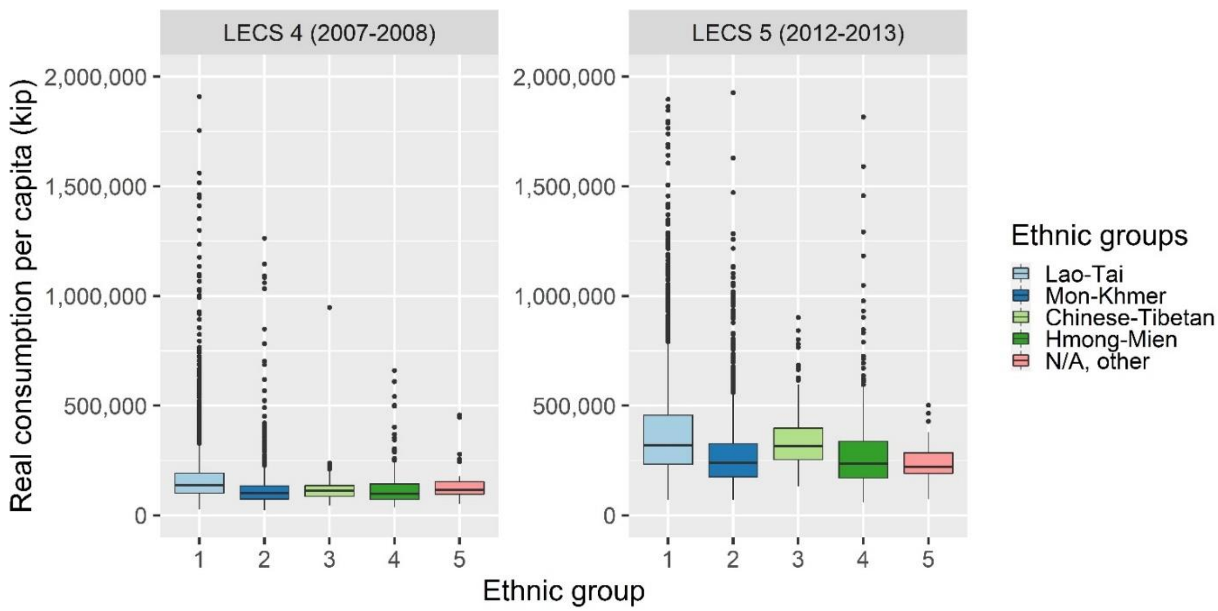


Figure A3.3: Real consumption per capita by ethnic minority group, 2007/08-2012/13. Source: Authors' estimates using the LECS4 and LECS5 Household Survey Data. Outliers over 2,000,000 kip removed for visualization purposes

Table A3.1 illustrates the Gini decomposition by ethnic group for rural areas. Only small changes in the Gini coefficients are observed from 2007/08-2012/13, although there are slight decreases for the Chinese Tibetan group, and slight increases for the Mon Mien. Our findings align with previous studies indicating that the Lao-Tai group exhibits a higher comparative level of inequality as compared to the Mon-Khmer and the Chinese Tibetan groups - the latter groups tend to have higher poverty rates (Epprecht et al. 2008). Between-group inequality is higher for this set of groups as compared to the rubber groups outlined in Table 6 in the main text, but does decrease slightly from 2007/08-2012/13. Overall, within-group inequality is the largest driver of overall inequality, echoing previous research (Warr et al. 2018).

Our results differ slightly from a previous study assessing inequality changes by ethnic group from LECS 3 (2002/2003) to LECS 5, which found uniform increases in the Gini coefficient across ethnic groups (Warr, Rasphone, and Menon 2015). However, our LECS sample is slightly different, due to the omittance of urban villages; additionally, we assess a shorter-term trend from LECS 4 to LECS 5. These divergences in our results make sense, given that urban-rural differences are one of the most important drivers of inequality in the country (Warr, Rasphone, and Menon 2015, Epprecht et al. 2008).

Table A3.1: Gini index decomposition for per capita consumption by ethnic groups (rural areas), LECS 4 (2007/08) and LECS 5 (2012/13)

Ethnic group	Population share LECS 4 (2007/08)	Gini coefficient	Population share LECS 5 (2012/13)	Gini coefficient	Change
Group 1: Lao-Tai	0.627	0.331	0.648	0.332	0.30%
Group 2: Mon-Khmer	0.249	0.318	0.251	0.319	0.31%
Group 3: Chinese Tibetan	0.025	0.235	0.035	0.211	-10.21%
Group 4: Mon Mien	0.091	0.312	0.057	0.329	5.45%
Group 5: Other	0.008	0.261	0.008	0.209	-19.92%
Within-group inequality, full sample		0.164		0.17	3.66%
Between-group inequality, full sample		0.078		0.063	-19.23%
Overlap inequality, full sample		0.094		0.1	6.38%
Gini coefficient, full sample	100%	0.336	100%	0.333	-0.89%
Source: Authors’ estimates using the LECS4 and LECS5 Household Survey Data.					

Table A3.2: Z-K Index for ethnic groups (rural areas only, alpha=0)

	LECS 4 (2007/08)	LECS 5 (2012/13)	Absolute Change
Ethnic group	**0.077**	**0.055**	-0.023
Between group inequality	0.013	0.010	-0.004
Within group inequality	0.174	0.174	0.000
Source: Authors’ estimates using the LECS4 and LECS5 Household Survey Data.			

We observe a slight decrease in polarization (as measured by the Z-K Index) when the consumption data is grouped by ethnic group, with reductions in between-group inequality. The Z-K values are higher when expenditures are grouped by ethnic composition as compared to rubber and region (see A3.4), which makes sense based on the importance of ethnic composition for inequality.

A3.3 Descriptive statistics of rubber groups

When households are grouped by rubber production and region, there are a few distinctive differences between rubber and nonrubber groups. These include differences in ethnic composition (Table A3.4). There are only minor differences in mean elevation between the rubber and nonrubber groups; notably, Group 1 has a lower mean elevation (600 m.a.s.l) than Group 2 (749 m.a.s.l., see Table A3.3). There are also differences in land access and ownership among rubber groups (Tables A3.5 and A3.6).

Table A3.3: Mean elevation of rubber groups

	Mean elevation (m.a.s.l)
Group 1: Northern Laos, with rubber	600
Group 2: Northern Laos, without rubber	749
Group 3: Central Laos, with rubber	364
Group 4: Central Laos, without rubber	403
Group 5: Southern Laos, with rubber	351
Group 6: Southern Laos, without rubber	362
Source: Authors' estimates using the LECS4 and LECS5 Household Survey Data.	

Table A3.4: Ethnic composition of rubber and nonrubber groups, LECS 4 (2007/08)

	Lao-Tai	Mon-Khmer	Chinese-Tibetan	Mon-Mien	NA/Other
Group 1: Northern Laos, with rubber	56.2%	3.8%	20.0%	20.0%	0.0%
Group 2: Northern Laos, without rubber	47.0%	31.0%	9.4%	12.4%	0.1%
Group 3: Central Laos, with rubber	59.7%	29.2%	0.0%	11.1%	0.0%
Group 4: Central Laos, without rubber	73.9%	15.8%	0.0%	9.2%	1.1%
Group 5: Southern Laos, with rubber	25.0%	50.0%	0.0%	0.0%	25.0%

Group 6: Southern Laos, without rubber	58.4%	41.6%	0.0%	0.0%	0.0%
Source: Authors' estimates using the LECS4 and LECS5 Household Survey Data.					

Group 1 has more Chinese-Tibetan households than Group 2. Groups 5 and 6 have very different ethnic compositions, Group 5 has a majority Mon-Khmer population (50%) followed by other ethnic minorities (25%), while Group 6 is primarily Lao-Tai (58%) followed by Mon-Khmer (42%).

Table A3.5: Land ownership of rubber and nonrubber groups, LECS 4 (2007/08) – LECS 5 (2012/13)

	Proportion owning land (2007/08)	Proportion owning land (2012/13)
Group 1: Northern Laos, with rubber	.94	.91
Group 2: Northern Laos, without rubber	.98	.96
Group 3: Central Laos, with rubber	.97	.95
Group 4: Central Laos, without rubber	.95	.95
Group 5: Southern Laos, with rubber	1	.98
Group 6: Southern Laos, without rubber	.94	.85
Source: Authors' estimates using the LECS4 and LECS5 Household Survey Data.		

Table A3.6: Land area ownership of rubber and nonrubber groups (hectares), LECS 4 (2007/08) – LECS 5 2012/13)

	Mean area owned (2007/08)	Mean area owned (2012/13)
Group 1: Northern Laos, with rubber	2.23	2.58
Group 2: Northern Laos, without rubber	2.21	1.99
Group 3: Central Laos, with rubber	1.80	1.76
Group 4: Central Laos, without rubber	2.23	2.25
Group 5: Southern Laos, with rubber	2.32	1.62
Group 6: Southern Laos, without rubber	2.44	2.15
Source: Authors' estimates using the LECS4 and LECS5 Household Survey Data.		

A3.4 Zhang-Kanbur (Z-K) Index Analysis

A3.4.1 Z-K Index Method

The Z-K index is a polarization measure reflecting the ratio of between-group inequality over within-group inequality, derived from the generalized entropy (GE) index (Zhang and Kanbur 2001). With lower within-group differences, and higher between-group differences, higher polarization is observed (Ibid). The Z-K index enables investigation of polarization by exogenously defined social groups, for example by ethnicity or geographic location (Indra et al. 2018, Antonio Duro 2010). Values of the index can help explain the importance of an exogenously defined group's contribution to polarization (Indra et al. 2018).

The formula for the Z-K index is defined below (3), based on the GE index of inequality (2) (Zhang and Kanbur 2001, Indra et al. 2018). The first term is equal to within group inequality and the second term is the between group inequality:

$$I(y) = \sum_{g=1}^{G} w_g I_g + I(\mu_1 e_1, \ldots \mu_K e_K) \qquad (2)$$

where:

$$w_g = f_g \left(\mu_g / \mu\right)^C;\ C \neq 0{,}1$$
$$f_g\left(\mu_g / \mu\right);\ C = 1$$
$$f_g\ ;\ C = 0$$

I_g is equal to inequality in the group g, μ_g is the mean of the group, and e_g is a vector of 1's of length n_g, where n_g is the population of the group g.

The index can then be expressed as:

$$P_{zk} = \frac{I(\mu_1 e_1 \ldots, \mu_k e_k)}{\Sigma_g^k = 1\ wgIg}$$

or

ZK = *between group inequality / within group inequality* (3)

We calculate the Z-K index based on the generalized entropy indices using the *IC2* package's "decompGEI" function in R (Plat 2012, R Core Team 2013).

A3.4.2 Z-K Index Results

We calculate the Z-K index for a variety of groupings to assess potential drivers of polarization. Overall, the Z-K index for the rubber groups overall remains stable from 2007/08-2012/13 (Table A3.7).

When we take a detailed look at within-group inequality for rubber groups as defined by the GE index (Table A3.9), we see that both the with and without-rubber groups in the North have experienced increasing polarization due to decreases in the within-group inequality; in the South, the reverse trend has occurred, and both with and without-rubber groups have experienced decreased polarization due to increases in within-group inequality.

These Z-K index results suggest that polarization driven by rubber production specifically did not change drastically over the study period, although slight differences occurred due to decreases in the within-group inequality in the North and increases in the within-group inequality in the Central and Southern regions. Between and within-group inequality values can be influenced by differences in subgroup sizes, and the number of subgroups (Epprecht et al. 2008), and the relatively small size of our with-rubber groups may drive the low Z-K values. The low Z-K index values for rubber may also reflect high overlap inequality among the groups.

For comparative purposes we also conducted Z-K analysis based on ethnic composition, finding higher Z-K values for those grouping; this suggests that ethnic minority composition is a larger driver of polarization.

Table A3.7: Z-K indices for rubber and region (rural areas only, alpha=0)

	LECS 4 (2007/08)	**LECS 5 (2012/13)**	**Absolute change**
Rubber	**0.013**	**0.013**	0.000
Between group inequality	0.002	0.002	0.000
Within group inequality	0.185	0.181	-0.004
Region	**0.009**	**0.005**	-0.004
Between group inequality	0.002	0.001	-0.001
Within group inequality	0.186	0.183	-0.003
Source: Authors’ estimates using the LECS4 and LECS5 Household Survey Data.			

Table A3.8: Z-K index-detailed look at regions, 2007/08-2012/13 (alpha=0)

	LECS 4 (between group inequality =	**LECS 5 (between group**	**Absolute change**

	.002)	inequality=.001)	
North	**0.008**	**0.006**	-0.002
Within group inequality	*0.223*	*0.160*	-0.064
Central	**0.011**	**0.006**	-0.005
Within group inequality	*0.160*	*0.166*	0.006
South	**0.010**	0.004	-0.006
Within group inequality	*0.177*	*0.239*	0.062
Source: Authors' estimates using the LECS4 and LECS5 Household Survey Data.			

Table A3.9: Z-K index-detailed look at rubber, 2007/08-2012/13 (alpha=0)

	LECS 4, between group inequality=.002	**LECS 5, between group inequality=.002**	**Absolute change**
Group 1	**0.020**	**0.024**	0.004
Within group inequality	*0.119*	*0.095*	*-0.024*
Group 2	**0.010**	**0.014**	0.004
Within group inequality	*0.227*	*0.162*	*-0.065*
Group 3	**0.016**	**0.014**	-0.002
Within group inequality	*0.147*	*0.165*	*0.018*
Group 4	**0.015**	**0.014**	-0.001
Within group inequality	*0.162*	*0.164*	*0.002*
Group 5	**0.031**	**0.010**	-0.020
Within group inequality	*0.077*	*0.223*	*0.146*
Group 6	**0.013**	**0.010**	-0.004
Within group inequality	*0.181*	*0.239*	*0.058*
Source: Authors' estimates using the LECS4 and LECS5 Household Survey Data.			

A3.5 Bibliography

Antonio Duro, Juan. 2010. "Decomposing International Polarization of per Capita CO2 Emissions." Energy Policy, Energy Efficiency Policies and Strategies with regular papers., 38 (11): 6529–33. https://doi.org/10.1016/j.enpol.2010.06.017.

Epprecht, Michael, Nicolas Minot, Reno Dewina, Peter Messerli, and Andreas Heinimann. 2008. "The Geography of Poverty and Inequality in the Lao PDR." https://doi.org/10.7892/BORIS.34330.

Indra, Indra, Suahasil Nazara, Djoni Hartono, and Sudarno Sumarto. 2018. "Expenditure Inequality and Polarization in Indonesia, 2002-2012." *International Journal of Social Economics* 45 (10): 1469–86. https://doi.org/10.1108/IJSE-02-2017-0051.

Plat, Didier. 2012. Inequality and Concentration Indices and Curves. R Package Version 1.0-1. R. https://CRAN.R-project.org/package=IC2.

R Core Team. 2013. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-project.org/

Warr, Peter, Sitthiroth Rasphone, and Jayant Menon. 2015. *Two Decades of Rising Inequality and Declining Poverty in the Lao People's Democratic Republic*. Asian Development Bank. https://www.adb.org/publications/two-decades-rising-inequality-declining-poverty-lao-pdr.

Warr, Peter, Sitthiroth Rasphone, and Jayant Menon. 2018. "Two Decades of Declining Poverty but Rising Inequality in Laos." *Asian Economic Journal* 32 (2): 165–85. https://doi.org/10.1111/asej.12147.

Zhang, X., and R. Kanbur. 2001. "What Difference Do Polarisation Measures Make? An Application to China." The Journal of Development Studies 37 (3): 85–98. https://doi.org/10.1080/00220380412331321981.