THE ROLE OF SMALL-SCALE FARMS IN THE GLOBAL FOOD SYSTEM

by

Vincent Ricciardi

M.Sc. Pennsylvania State University, 2014

B.A. Kalamazoo College, 2009

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF

THE REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

in

THE FACULTY OF GRADUATE AND POSTDOCTORAL STUDIES

(Resource Management & Environmental Studies)

THE UNIVERSITY OF BRITISH COLUMBIA

(Vancouver)

September 2019

The following individuals certify that they have read, and recommend to the Faculty of Graduate and Postdoctoral Studies for acceptance, the dissertation entitled:

The role of small-scale farmers in the global food system

submitted by Vincent Ricciardi in partial fulfillment of the requirements for

the degree of Doctor of Philosophy

in Resource Management & Environmental Studies

Examining Committee:

Navin Ramankutty

Supervisor

Hannah Wittman

Supervisory Committee Member

Milind Kandlikar

Supervisory Committee Member

Rashid Sumaila

University Examiner

Jeff Sayer

University Examiner

Abstract

Farms are becoming larger in high-income countries and smaller and more fragmented in low-income countries. This farm size transition has motivated recent international policy calls to combat inequities in the global food system and to promote more environmentally friendly production practices. This dissertation seeks to understand the impact of this transition on global food production, on the environment, and on small-scale farmers' livelihoods. Each chapter aims to examine the underlying assumptions of these policy calls by testing the relationship between small-scale farms and key socio-ecological outcomes. Through creating a harmonized dataset across 55 countries, Chapter 2 estimates that 30-34% of the world's food is produced by farms under 2 hectares in size and smaller farms producer a greater diversity of crop species than larger farms. Chapter 3's meta-analysis synthesizes the past 50 years of empirical evidence on the relationships between farm size and several socio-ecological outcomes of farming systems (i.e., yield, biodiversity, resource-use efficiency, greenhouse gas emissions, and profitability). We found that smaller farms have higher yields and promote more non-crop biodiversity (across species at the farm and landscape scales) than larger farms. We found no relationship between farm size and resource-use efficiency, but smaller farms had a non-significant trend to lower greenhouse gas emissions per crop output than larger farms. Chapter 4 builds on this meta-analysis to investigate the relationship between farm size, productivity, and income using harmonized national sample surveys from 34 countries across the Global South. Our results highlight that while smaller farms are more productive than larger farms, they have lower per capita incomes.

Critically, we also found that the current internationally agreed upon target of doubling small-scale farmer incomes is not aggressive enough to transition them out of poverty. Chapter 5 outlines pathways for the international community to identify and monitor small-scale food producers, which is a central hurdle to ensure governments keep their agreed upon commitments to support small-scale farmers. This dissertation makes several empirical contributions to the literature on the contribution of small farms to the global food system, and can help inform policy initiatives aimed to support small-scale farmers.

Lay Summary

The majority of the world's farms are smaller than two soccer fields. Despite representing the majority of all farms, small-scale farmers are amongst the most impoverished and food insecure populations in the world. This dissertation seeks to understand relationships between farm size, crop production, farmer livelihoods, and the environment. The results suggest that smaller farms grow a large amount of the world's food, promote crop and non-crop biodiversity, and are more productive than larger farms. Yet, we found that many smaller farm households still live below their national poverty lines. These results can help inform current international initiatives aimed to support small-scale farmers, such as the Sustainable Development Goals, by refining development targets and by establishing pathways to monitor the role of small-scale farms in the global food system.

Preface

This is a manuscript-based dissertation, where Chapters 2-5 were written as independent articles. Since each article was prepared for a different publication outlet, there are some overlap contained in the introductions. The overall structure and goals of this dissertation were codesigned with my advisor, Navin Ramankutty, and committee members, Hannah Wittman and Milind Kandlikar. While I am the lead author and analyst on all chapters, I worked collaboratively with several coauthor teams.

A version of Chapter 2 was published as: Ricciardi, V., Ramankutty, N., Mehrabi, Z., Jarvis, L. and Chookolingo, B., 2018. How much of the world's food do smallholders produce? *Global Food Security,* 17, pp.64-72. A large portion of the supplemental material for Chapter 2 (Appendix A.2 and A.3) was published as: Ricciardi, V., Ramankutty, N., Mehrabi, Z., Jarvis, L. and Chookolingo, B., 2018. An open-access dataset of crop production by farm size from agricultural censuses and surveys. *Data in Brief*, 19, pp.1970-1988. For these outputs, I harmonized the survey data with assistance from a small team of research assistants that I trained and managed, I conducted the analysis, wrote the first draft of the article, and incorporated co-author edits on subsequent revisions. Co-authors assisted in a combination of data processing, analysis guidance, and/or editing.

A version of Chapter 3 is in the process of being submitted to a peer-reviewed journal with Navin Ramankutty, Hannah Wittman, Zia Mehrabi, and Dana James as coauthors. For this meta-analysis, I designed the protocol, I was the primary data collector, I conducted the analysis, wrote the first draft of the article, and incorporated co-author edits on subsequent revisions. Co-authors assisted in a combination of data processing, analysis guidance, and/or editing.

Chapter 4 has been prepared for peer-reviewed publication with Navin Ramankutty and Zia Mehrabi. This chapter was in collaboration with the Food and Agricultural Organization (FAO), in which my analysis was being used to both detect data anomalies in their new dataset and as a demonstration for their dataset's utility. I conducted the analysis, wrote the first draft of the article, and incorporated co-author edits on subsequent revisions. Co-authors assisted in analysis guidance and editing.

A version of Chapter 5 has been prepared to be submitted as a policy paper to an international development outlet. This chapter was written in collaboration with Navin Ramankutty, Hannah Wittman, Zia Mehrabi, and Balsher Sidhu. I conducted the analysis, wrote the first draft of the article, and incorporated co-author edits on subsequent revisions. Co-authors assisted in analysis guidance and editing.

Since no primary data was collected, ethics approval was not required for this dissertation.

Table of Contents

List of Tables

List of Figures

Acknowledgements

I feel extremely lucky to be able to pursue a doctorate degree, nonetheless, with such an amazing group of people. I am tremendously grateful to my advisor, Navin Ramankutty, for all of his guidance and support throughout the dissertation process. He has been a phenomenal advisor, has truly pushed my analytical thinking, and has taught me to approach science with a humble mind. A special thanks to Hannah Wittman for always challenging my rhetoric and framing -- these lessons have been essential for me to ground my findings in pressing policy conversations. To Zia Mehrabi, for all of his collaborative brainstorming on statistics and framing. This dissertation would not have been as fun without my excellent and supportive lab mates -- thank you for putting up with my terrible jokes.

There are countless people in our department, university, and research communities that have collectively fostered a safe space for me to admit my knowledge gaps and to share their insights openly. A broad thanks goes out to everyone in our department and land systems science community for the excellent and cross-disciplinary conversations. I am especially grateful to Piero Conforti and his team at the Food and Agricultural Organization (FAO) for sharing their dataset that underlies Chapter 4.

Without my amazing family I would have never had the idea nor chance to dedicate these years to pure curiosity. My parents and siblings have instilled in me a work ethic and sense of humor that were essential to this process. Page, my partner, has been especially amazing throughout this degree. She has helped edit and think through every stage of this dissertation -- she has also persevered when I often ranted about farm size only to go silent for 20 minutes, then picked up where I left off. I have no idea how she followed all these conversations and she did so with incredible composure and insight.

This research was funded through a University of British Columbia 4 Year Doctoral Fellowship & Social Sciences and Humanities Research Council (SSHRC) Insight Grant (#435-2016-0154). Without this financial support, this dissertation would not have been possible. Finally, I would like to thank Ralph.

To my grandfather, who taught me to love food and laugh often.

.

Chapter 1: Introduction

1.1 Overview

The majority of the world's farms are small -- of the 570 million farms in the world, 84% are less than two hectares (ha) in size and constitute 12% of farmland area (1). Many scholars claim that smaller farms are more productive, resource-use efficient, and environmentally friendly than larger farms (2–4), and yet, smallholder (a synonym for small-scale farmers) livelihoods are facing pressure from low prices in global markets and production losses from climate-change induced extreme weather (5). Smallholders are amongst the most impoverished and food insecure populations in the world (6, 7).

An added stress to smallholder livelihoods is that the scale of agricultural production is rapidly changing. In general, farms are becoming larger and more industrialized in high-income countries, and smaller and more fragmented in low-income countries (Figure 1) (1). The primary factors mediating changes in farm size are economic development, land consolidation and redistribution policies, colonial legacies, traditional land inheritance systems, and Large-Scale Land Acquisitions (LSLA) (8, 9). Depending on how a country develops its non-agricultural sectors, farmers can be pulled out of agriculture when there are other employment opportunities -- a process that often leads to larger farms. However, farmers can also be pushed out of farming when land consolidation policies are enacted, and pushed into poverty and food insecurity if such policies are enacted before the economy is able to absorb the resulting agricultural work-force into non-agricultural employment (9). Thus, changes in farm size can be a driver of food insecurity and poverty, as well as a response to broader economic growth and improved food security. Across socio-political and environmental contexts, the cumulative impacts of farm size transitions remain largely understudied for both farmers and wider society.

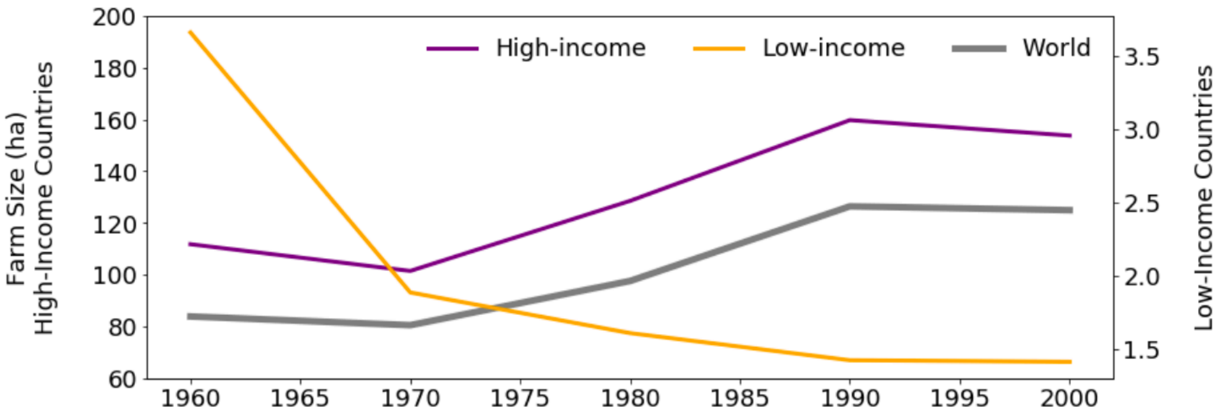

Figure 1: Globally, the average farm size is increasing (left axis). Farms in high-income countries are getting larger (left axis) and farms in low income countries are getting smaller (right axis). Data from Lowder et al. (1).

There has been growing concern about the need to promote economic growth that is inclusive of small-scale farmers. From the 1990s through 2010, there has been a surge of national level policies targeting improved incomes and productivity of small-scale farms (Figure 2). Recently, international efforts by civil society organizations have shifted to coordinate donor resources explicitly to include smallholders into sustainable development agendas. The most widely recognized of these international efforts is Goal 2 of the UN Sustainable Development Goals (SDGs) that aims to end hunger and achieve food security through sustainable agriculture; a key target (SDG 2.3) is by “2030, [to] double the agricultural productivity and incomes of small-scale food producers, in particular women, indigenous peoples, family farmers, pastoralists, and fishers” (10). Several other international efforts in support of smallholders preceded SDG 2.3. In 2014, the “International Year of the Family Farm,” the United Nations (UN) and other food security agencies reinforced the need to support for family farmers (11), a term often interchanged or analogous with smallholders (12). The COP21 agreement (the 2015 UN Conference of Parties on Climate Change) resulted in 179 country commitments that include bolstering smallholder adaptive capacity to climate change. Apart from these sustainable development initiatives, consumers in higher-income countries are increasingly concerned about health, farmer livelihoods, and the decline of diverse and traditional foods; these trends have led to an increased willingness-to-pay for products with organic and/or local labels, which are often associated with smaller farms (13–17). Despite progress in including smallholders and consumer interest in smaller farms in international policy agendas, there is scant empirical data on smallholder farms and their role in the global food system.

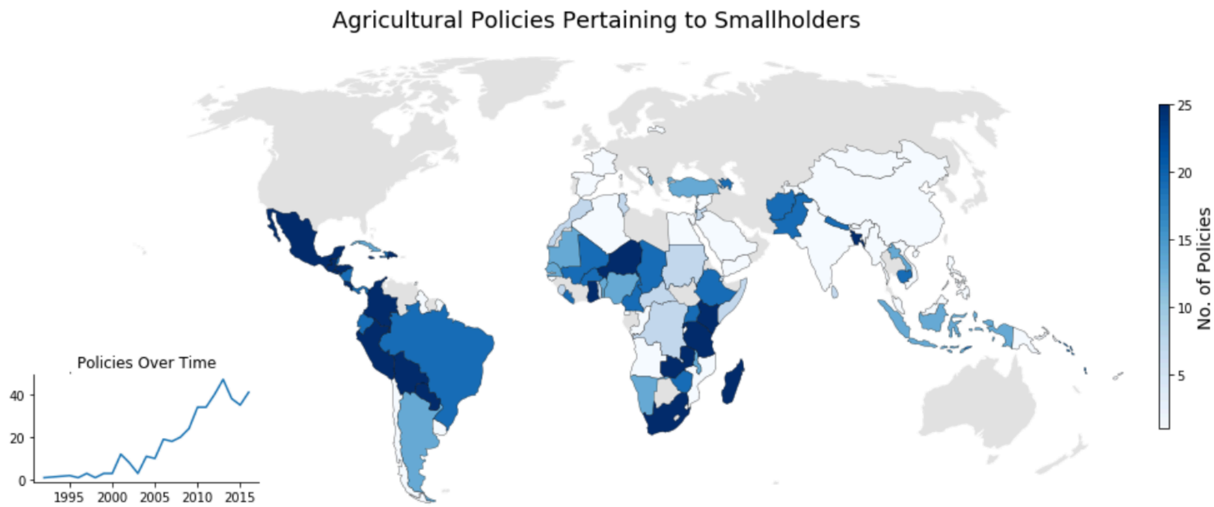

Figure 2: Data from the FAOLEX database, which contains over 132,000 policies, legislation, regulation, international agreements, and constitutions from around the world from 1990 to 2015. A search for the keyword "smallholders" revealed a cumulative 443 national policies, legislations, regulations, and/or constitutions for 127 countries (out of 196 total).

To support sustainable transitions in farm size, promote international policy calls, and hold governments accountable in their agreements for inclusive economic development, the research and development community needs a better understanding of the role of small-scale farms in the global food system. A central challenge to conducting global research on small-scale farms has been data availability and the need to study complex drivers and outcomes of smallholders' food insecurity. For instance, while there are many national sample surveys and agricultural censuses available, they use diverse definitions of farm size (e.g., operational area, harvested area, etc.), contain different crop species and using colloquial names, and are conducted at different scales of statistical representativeness. Similarly, the literature on small-scale farms is challenging to assess through systematic review and meta-analysis techniques because of the inconsistent definitions for smallholders, the range of analytic methods used across disciplines, the diversity of types of small-scale farms, and the range of socio-political and environmental contexts that case studies cover.

Despite these challenges, global-scale research on small-scale farms has made several major contributions since 2015 (1, 18–21). Empirical studies have detailed global trends in farm size (1), the amount of agricultural land operated by family farms (18), and the types and quantity of crops produced by smaller farms (19, 20). Systematic reviews and meta-analyses on small-scale farms have examined the relationships between production and dietary diversity in smallholder households (22) and environmental impacts across farm sizes (23). These studies represent a transition in empirical understandings of smallholders at national or regional scales to the importance of smallholder systems at the global-scale for: food production, farmer food security, and as possible pathways toward sustainable agriculture. While these themes have been championed by certain academic and advocacy groups for many years, large-scale empirical analyses to test these claims has been missing. Global scale analysis allows for cross-context comparisons to highlight the environmental, poverty, and production relationships and the types of policy solutions and/or farm-level innovations that can be applied from one context to another.

1.2 Objective and research questions (RQs)

The broad objective of this dissertation is to understand the role of small-scale farms in the global food system. Building upon the limited literature of global data analyses and synthesis studies, I seek to understand centrally debated topics regarding small-scale farms. To contextualize small-scale farms, I compare small-scale farms with larger farms to identify how agricultural production, yield, environmental impacts, and income are affected by changes in farm size. More specifically, the following research questions are examined (see Table 1 for detailed sub-questions):

1. How much food do small-scale farmers produce globally? Do certain farm sizes produce more food, animal feed, seed, or other types of crop uses? Do smaller farms grow a greater diversity of crop species than larger farms?

2. How do yields, biodiversity, resource-use efficiency, greenhouse gas (GHG) emissions, and profitability vary by farm size? How do social, political, and environmental contexts mitigate these relationships?

3. What is the relationship between profit per ha, income per person, and farm size?

4. Is farm size a good indicator of farmer poverty or do other dimensions of small-scale farms (e.g., farm size, economic size, family farms, subsistence production, etc.) offer better proxies for farmer poverty? Which combination of definitions can operationally identify impoverished farmers for SDG 2.3?

5. Is the SDG 2.3 target of doubling the incomes of small-scale farmers enough to raise them over their national poverty lines?

Table 1: Overview of dissertation's themes, key socio-ecological outcome variables, main and sub-research questions, and which chapter answers each question. A detailed literature review and hypotheses are given in each corresponding chapter. The conclusion presents fidings for each question in this table.

Theme	Variable	RQ	Sub-Research Question	Chapter
Production	Food produced	1	How much of the world's food is produced by smallholders?	2
	Non-food produced	1	Within each farm size class, how much of the crop production is allocated to different uses?	2
	Yield	2	Do smaller farms have higher yields than larger farms?	3
Environment	Crop diversity	1,2	Do smaller farms grow a greater diversity of crops than larger farms?	2,3
	Non-crop diversity	2	Do smaller farms promote more on-farm and landscape level non-crop biodiversity than larger farms?	3
	Resource-use efficiency	2	Do smaller farms have greater resource-use efficiency than larger farms?	3
	Greenhouse gas emissions	2	Do smaller farms have lower greenhouse gas emissions per unit of crop output than larger farms?	3
Socio-Economic	Economic productivity (profit per ha)	3	Are smaller farms less productive (profit per ha) than larger farms?	4
	Income (profit per capita)	3	Do smaller farms have lower incomes (profit per capita) than larger farms?	4
Policy	Small-scale farm definition	4	Is farm size a good indicator of farmer poverty compared to other small-scale farm definitions?	4
	Doubling incomes as an SDG target	4	Is the SDG 2.3 target of doubling the incomes of small-scale farmers enough to raise them over their national poverty lines?	4

1.3 Defining small-scale farms

Small-scale farms have a diversity of definitions that range from the spatial operating size of a farm to the environmental, socio-political, and economic aspects of a farm. To address these range of definitions, I first define small-scale farms in terms of their spatial operating size. Then, I later empirically examine how spatial farm size compares to other commonly used dimensions of small-scale farms (e.g., family farm, economic size, and country relative farm size).

My initial focus on the spatial size of a farm addresses three motivations. First, operating area has been a critical dimension for land consolidation and land reform policies (9), both of which are common rallying points for smallholder advocates (24, 25), and has been conventionally used in national sample surveys and empirical literature (1, 26). To answer RQ1 and RQ2, available household surveys are harmonized and a meta-analysis is conducted, which reflects the past 50 years of defining small-scale farms in spatial terms. In addition, recent advances in Earth observation has enabled easier classification of field size, which is currently being linked to farm size (27). While farmers operating under 2 ha have been defined as smallholders conventionally, to answer RQ1, I present both widely used breakpoints for small-scale farms and a continuum of farm sizes for the reader to use their own preferred cut-off points. The meta-analysis used to answer RQ2 examines how changes in farm size -- opposed to cut-off points -- may affect different socio-ecological outcomes. Second, consumers in higher income countries are increasingly concerned about health, farmer livelihoods, and the decline of diverse and traditional foods, leading to an increased willingness-to-pay for products with organic and/or local labels, which are often associated with smaller farms (13–17).

Finally, spatial farm size is one of the two dimensions of small-scale farms that have been proposed by the FAO to operationalize the SDG 2.3 target. To monitor SDG 2.3, a combined country-relevant definition of smallholders would identify a country's smallest 40% farms (as defined by spatial size) and farms within the bottom 40% of economic size (as defined by agricultural revenue) (28). Since there has been little cross-regional and global scale work to explore the relationships between farm size and intended outcomes of the SDGs (e.g., poverty alleviation, decreased environmental impacts, and increased productivity), each dimension of this definition needs to be examined. As one key dimension of the SDG 2.3 operational definition, farm size's connection to ongoing land-use policies, available data, historical usage, and consumer interest, and limited analysis across regional contexts and disciplinary boundaries suggests a pressing literature gap.

In this dissertation, I also use non-spatial definitions of farm size. Spatial definitions of farm size have been criticized for being arbitrary, lacking a country relevant perspective (e.g., by the 2 ha definition, Brazil would have 20% of its farms classified as smallholders, while India would have ~80% of its farms be classified as smallholders), and not accounting for other key dimensions of impoverished farmers, such as a farms' economic size (e.g., annual revenue), labor dimensions (e.g., family labor, female labor, etc.), and market orientation (e.g., level of subsistence) (1, 12). In addition, global statistics on the role of small-scale farms in the food system are obfuscated by differing definitions, such as "small farms" (1, 19, 20, 29), "peasant farms" (30), and "family farms" (11, 18).

To answer RQ4, I examine how spatial farm size relates to these other key dimensions associated with small-scale farms and how each dimension relates to poverty and economic productivity -- the key SDG 2.3 targets. While past studies have relied on policy grounded, yet theoretical arguments that are in support of various definitions, my research provides novel empirical tests for these different definitions of small-scale farms. My perspective fulfills a key literature gap to support these theoretical arguments with newly available and harmonized data from 34 countries across the low and middle-income countries (LMIC).

1.4 Structure of dissertation

Chapter 2 examines how much food small-scale farms produce globally through harmonizing a dataset of nationally representative household surveys and agricultural census data. The harmonized dataset was made open-access with a separate data article detailing the dataset construction and testing key assumptions in its construction (Appendix A). Chapter 2 is the study to directly evaluate the relationship between farm size, crop types, and crop diversity across a large range of farm sizes and geographic regions, and to assess how this diversity influences the amount of macro-nutrients available from crops.

Chapter 3 synthesizes 50 years of literature to identify different socio-ecological outcomes of farm size, such as the relationship between farm size and yield, crop diversity, resource use efficiency, GHG emissions, and profitability. This effort highlights the need to synthesize the scientific knowledge of small-scale farms to understand under what socio-political and environmental conditions are required for certain production, environmental, and economic outcomes of farm size to arise. This study is among few efforts that bridge siloed academic communities that research different aspects of small-scale agriculture (e.g., economists and ecologists). Amongst the relationships examined in Chapter 3, smaller farms consistently had higher yields than larger farms across a large array of geographic contexts. However, the limited number of studies examining profitability of small-scale farms compared to larger farms left a greater question unanswered: is the higher yield of small-scale farmers enough to raise them out of poverty?

Chapter 4 seeks to understand the relationship between smallholders, economic productivity, and income through a partnership with the UN Food and Agricultural Organization (FAO), which has harmonized national sample surveys from 34 countries and will release this dataset in late 2019. This study looks at the difference between profit per ha (productivity) and profit per person living in an agricultural household (income). This study then goes beyond spatial farm size definitions to explore how other commonly used dimensions of smallholders (e.g., family labor, market-orientation, country relative farm and economic size) map onto poverty.

Chapters 5 and 6 conclude the dissertation. Chapter 5 summarizes the key findings of Chapter 4 into a policy brief discussing SDG 2.3's small-scale farm definition, and provides an empirical example of future data harmonization work that can be used to monitor countries' progress towards SDG 2.3. This policy brief suggests improvements to the UN proposed binary definition of small-scale farmers. It then details the present landscape of data that can help contextualize where different groups of small-scale farmers live to target needed resources (e.g., what percent of small-scale farmers live in water-scarce areas?). Chapter 6 concludes the dissertation by summarizing the key findings.

1.5 Limitations

This dissertation has several limitations including the type of data required for a global analysis. These limitations should be taken into consideration when interpreting my results. Each chapter contains its own separate discussion on the limitations of the study and provides sensitivity tests to estimate the effects of each assumption on the results. The following limitations are the most salient throughout the dissertation.

By seeking a global perspective, I needed to harmonize datasets from disparate data sources; a process that required assumptions to combine surveys with different variables. In Chapter 2, I led a team to harmonize a dataset from national sample surveys and agricultural censuses across 55 countries. In Chapter 4, I collaborated with the FAO to refine a harmonized dataset from national sample surveys across 37 countries. For both datasets, I examined key variables for data anomalies within and across countries, performed sensitivity tests to ensure my conclusions were not due to faulty data, and employed methods that deliberately accounted for differences in survey quality, design, and sampling strategies. In Chapter 2, I had full control of the data harmonization process. I publicly released the dataset and all processing code as well as published a separate data article (Appendix A) to explain our assumptions and their likely effects on our conclusions. In Chapter 4, I relied upon the FAO's RuLIS dataset, whose team has produced technical documentation to explain their processing steps and their key assumptions. While I attempted to take precautions with constructing and analyzing these datasets, there are most likely issues from the surveys themselves (e.g., recall bias, and errors in sampling or data entry) as well as unintended consequences of assumptions that I was not aware.

The dataset in Chapter 2 represented a global convenience sample, with extrapolations to countries not included in the dataset. While we did not include several large agricultural countries due to data availability, such as China, we found our results comparable with two other studies that used geospatial data with more complete spatial representation. The dataset in Chapter 4 was a convenience sample across the Global South. While analysis of 37 countries is not "global," the cross-country comparative analysis gives insight onto global processes, and is representative of a diversity of countries with smallholder populations. While there were no other analyses that we could compare these results to, since the scale for this analysis was novel for the topic, we used methods that are widely acknowledged to provide generalizable insights to countries not included in the dataset (e.g., mixed effects models as outlined in Chapter 4) (137); in this chapter, we limited our general conclusions to other LMICs. For these reasons, I urge the reader to be critical of these results, while acknowledging that these datasets represent the current state of knowledge on small-scale farms at a global scale for Chapter 2, and across the Global South for Chapter 4.

In Chapter 3, I led a systematic review and meta-analysis of six different socio-ecological outcomes of cropping systems. Synthesis studies can provide insight into consensus and disagreements in the scientific literature, but the research questions, search strategy, and the methods need careful consideration. There were two central limitations to Chapter 3. First, our meta-analysis required empirical studies that directly compared farm size to one of our outcome variables. This narrow inclusion criteria allowed us to examine the trade-offs between changes in farm size and several socio-ecological dimensions. However, we were not able to assess non-quantitative studies or studies that did not examine changes in farm size. While these were beyond the scope of Chapter 3, we acknowledge that there is a vast literature that details additional outcomes of agricultural systems related to small-scale agriculture. For example, small-scale farms can promote cultural preservation and rural community cohesion (24, 115), but larger farms can be more suitable for precision agriculture, which may reduce environmental impacts (62). Another key limitation of Chapter 3 was regarding the availability of academic databases to which we had institutional access. We relied on Web of Science and Scopus for our literature search, which may have limited our ability to include all peer-reviewed studies that would have meet our inclusion criteria. To address these limitations, I detail our search strategy in Appendix B and plan to release the dataset and analysis code with the chapter's peer-reviewed publication.

In Chapter 4, I led a comparison of several definitions of small-scale farms to understand which definitions can be used to operationally monitor SDG 2.3. A key limitation of this study is that we were only able to test a limited number of definitions. We tested farm size (in hectare and country relevant terms), country relevant economic size, the reliance on family labor, and subsistence levels. While these five definitions are commonly discussed in the small-scale farm literature (see Chapter 4 for a detailed overview), we were unable to test other pertinent definitions with the available data (e.g., rainfed or remoteness).

To help overcome the above limitations, I am in the process of publicly releasing all the underlying datasets and analysis code. My hope is that by providing open-access datasets and analysis code built with freely available software, there will be more transparency in the derivation of global statistics on small-scale farmers. Also, other researchers may add to these datasets and analysis packages.

While acknowledging these limitations, this dissertation attempts to synthesize the current state of knowledge about small-scale farms in the global food system. The past 50 years have resulted in a wave of research and surveys about small-scale farms that provide rich and multi-faceted information. By leveraging these past research outputs, this global synthesis will enable a broader perspective across often-siloed academic disciplines and geographic regions. The aim of this thesis is to synthesize and build upon previous empirical and theoretical findings, as these are often locally derived and yet underpin the current wave of international support for small-scale farmers.

Chapter 2: How much of the world's food do smallholders produce?

2.1 Abstract

The widely reported claim that smallholders produce 70–80% of the world's food has been a linchpin of agricultural development policy despite limited empirical evidence. Recent empirical attempts to reinvestigate this number have lacked raw data on how much food smallholders produce, and have relied on model assumptions with unknown biases and with limited spatial and commodity coverage. We examine variations in crop production by farm size using a newly-compiled global sample of subnational level microdata and agricultural censuses covering more countries (n = 55) and crop types (n = 154) than assessed to date. We estimate that farms under 2 ha globally produce 28–31% of total crop production and 30–34% of food supply on 24% of gross agricultural area. Farms under 2 ha devote a greater proportion of their production to food, and account for greater crop diversity, while farms over 1000 ha have the greatest proportion of post-harvest loss.

2.2 Introduction

It has been widely reported that smallholder farmers (defined generally as being less than 2 ha) produce 70-80% of the world's food (11, 30, 31), are central to conserving crop diversity (2, 32, 33), produce more food crops than larger farms (34, 35), and yet are largely food insecure (36). These arguments have been a linchpin in recent agricultural development policy. For example, in

2014, the ‘International Year of the Family Farm’, the United Nations (UN) and other food security agencies reiterated these arguments to garner increased support for family farmers, who are predominantly smallholders (11). The COP21 agreement (the 2015 UN Conference of Parties on Climate Change) includes mitigation and adaptation commitments pertaining to agriculture from 179 countries that include the need to bolster smallholder adaptive capacity to climate change. Goal 2 of the UN Sustainable Development Goals (SDGs) aims to end hunger and achieve food security through sustainable agriculture; a key target (SDG 2.3) is by “2030, [to] double the agricultural productivity and incomes of small-scale food producers, in particular women, indigenous peoples, family farmers, pastoralists, and fishers” (37). Yet, despite progress in steering development policy towards smallholder farmers, there is scant empirical data on smallholder farms, and their role in the food system.

Key to enacting and monitoring progress on these international agreements and policies is a global baseline on the contribution of smallholders to global food production and security. However, the data underlying three widely reported claims on smallholder crop production remain non-transparent or contradictory. First, the source of various UN reports citing smallholder production is a communiqué from the ETC group (30), which suggests that “peasants” grow at least 70% of the world’s food; yet, the derivation of the estimate is obscure in this report. Second, the claim that smaller farms produce more food directly consumed by people, with larger industrialized farms producing more non-food crops, such as biofuels and animal feed (34, 35), has been brought into question by the observation that smaller farms have larger amounts of post-harvest loss due to lack of market and cold storage access (38, 39). Thirdly, while some authors argue that economies of scale are needed for farms to produce a diversity of crops (40), others suggest that larger farms face labor constraints that hamper mixed-cropping systems (41), so it is unknown if smaller farms produce a greater diversity of crop species than larger farms. In sum, our current understanding how much food smallholders produce, what kinds of food they produce, where their food is destined in the food system, and how much nutrition it contains, are all key knowledge gaps in global agricultural research.

The need to fill these knowledge gaps has been recently recognized by scientists (1, 18–20) (referred to as Graeub, Lowder, Herrero, and Samberg respectively hereafter). In 2016, a pair of studies evaluated the contribution of smallholders and family farms to global crop and food production. Lowder was the first to report on global farm size trends from 1960 to 2010 derived from 167 countries in the World Census of Agriculture (WCA). They found that small-farms (defined as being < 2 ha) constituted only 12% of the global available farmland, but represented 84% of all farms. Their study did not report on crop production, but their results implied that smallholders do not produce 70% global crops; it is unlikely they could produce this much food on 12% of available farmland, even if we assumed that small farms had higher yields and produced more food crops than larger farms. The second of these studies (Graeub) quantified the number and extent of family farms in the world and their production contributions. By using national family farm definitions, defining family farms based on farm size, or a combination thereof to represent regionally appropriate family farm definitions they estimated that ~98% of all farms globally are family farms, collectively managing 53% of all cropland, and meeting an estimated 36–114% of domestic caloric requirements for different countries. While Graeub's study highlighted the contribution of family farms, they also challenge the idea that all family farms are small farms. For example, farms in Brazil may be family owned but are large in size (while ~85% of farms in Brazil are family owned and cover ~25% of agricultural land, only 21% of farms are less than 2 ha in size and cover only 0.25% of the agricultural area). Together these two studies, quantified the global number of smallholders or family farmers, their cropping area, and detailed the differences between smallholders and family farms.

In 2017, two additional studies were published that tried to better estimate the proportion of food coming from smallholder farmers globally. Samberg estimated the contributions of smallholders in an analysis of 41 crops and 83 countries in smallholder dominant regions (Latin America, sub-Saharan Africa, and South and East Asia) that represent 35% of global cropland. They estimated that smallholders (which they defined as all administrative units with a "mean agricultural area" < 5 ha) produced 52.5% of food calories in their cross-regional sample. While, this study was a valuable step in mapping the geographic distribution of smallholders, using mean agricultural area within an administrative unit as an index of smallholder production is problematic because farm size distributions are highly skewed (e.g. Lowder). Following this, Herrero presented an analysis which modelled crop and livestock production, micro-nutrition production, and agricultural landscape diversity. Crop and animal data were related to farm size classes by combining crowd-sourced data on field sizes (21) with national farm size distributions (Lowder) as a proxy for per pixel production by farm size. They reported that farms < 50 ha produce 51–77% of commodities and nutrients in their sample of 41 crops, 7 livestock, 14 aquaculture and fish products, across 161 countries. They also estimated that ~20% of food calories globally come from farms < 2 ha, and highlighted the valuable micronutrient contribution of smallholders, with farms < 20 ha producing ~70% of the world's vitamin A. While both Samberg and Herrero provided clear steps forward in understanding the role of smallholders in the food system, and in particular Herrero covering both animal and crop products, they did not use direct measurements of crop production and/or area by farm size, compute diversity calculations based on these direct calculations of production and/or area, or report on the broader role of smallholders in the food system (e.g. how much of their food is wasted and destined to non-food crops).

To fill the gaps in directly linking farm size with crop production that previous studies were not able to examine, we compiled the first open source dataset to estimate crop production by farm size derived from actual farmer surveys containing crop-specific measurements of production or area that are cross-tabulated against each farm size class. Our dataset includes 154 crop types and covers 55 countries, which represents 51.1% of global agricultural area. We compare these direct estimates to those from the previous modeling studies (e.g., Herrero; Samberg). In addition, we provide global estimates of the type of production (i.e., food, feed, processing, seed, waste, and other) across farm sizes and within each farm size class, to understand if more production from small farms is wasted from storage and transportation, and if this cancels the larger losses to biofuels and animal feed grown on large farms. Finally, we evaluate how the type of crops grown, crop species diversity, and macro-nutrient production varies by farm size. Our study is the first to directly evaluate the relationship between farm size, crop types, and crop diversity across a large range of farm sizes and geographic regions, and to assess how this diversity influences the amount of macro-nutrients available from crops. Together, these results provide the most comprehensive empirically grounded estimates of crop production by farm size currently available.

2.3 Methods

2.3.1 Data compilation

We compiled a global convenience sample of datasets that directly measured crop production and/or area by farm size for 55 countries at either the national, or subnational level (for a total of 3410 national or subnational units; see Figure 3). Our dataset represents ~50% of the global agricultural area (see Appendix A for dataset construction and coverage). These datasets were either agricultural census data or nationally (or sub-nationally) representative sample surveys, aggregated by administrative unit (n = 34 countries) or available at the micro-level (e.g., anonymized individual household level records) (n = 21 countries; of which 18 were household surveys and 3 were censuses that captured both family and non-family farms). The median year of the data was from 2013, with the oldest datasets from 2001 and the newest from 2015. The database has 154 crops which we matched with commodity names outlined in the Food and Agricultural Organization's (FAO) statistical database (42) [FAOSTAT hereafter]. Where farm size and production were not cross-tabulated in the survey instrument (i.e. for 33 countries), we calculated production by farm size by first extracting either harvest area, cultivated area, crop area, or planted area to calculate farm size, and then converted area to production using FAOSTAT's national yield data. We tested the validity of this method, and found it to slightly underestimate production (full details of bias tests, inclusion criteria, variable descriptions, summary statistics, and per country statistics are given in Appendix A). When farm size data was not available for a country, but we had micro-level data, we used the sum of farm plot areas for a given household as a proxy for farm size. Internal validation of the use of micro-data to fill in data gaps was not possible with our data, because we did not have both micro-data and farm size metrics for any of our countries, but we think the impact of using aggregate plot area is likely to be negligible for our results, as this was only used on 4.8% of administrative units in our dataset. Finally, all crop production data was tallied per country and validated against available national level reports, and to the FAOSTAT crop production database, both of which are computed from aggregated crop area estimates. In total, our dataset captures 51.1% of global crop production and 52.9% of global cropland area. We harmonized the datasets to match the WCA farm size categories: 0 to 1 ha, 1 to 2 ha, 2 to 5 ha, 5 to 10 ha, 10 to 20 ha, 20 to 50 ha, 50 to 100 ha, 100 to 200 ha, 200 to 500 ha, 500 to 1000 ha, and above 1000 ha. While we recognize that per country definitions of smallholders may not fall within these farm size bins, the majority of the datasets included reported these farm size breaks. We report our estimates by each WCA farm size class and cumulatively to allow flexible definitions of smallholders that are consistent with past attempts to quantify the relationship between farm size and crop production. Future researchers may use the accompanying, open-access dataset to redefine smallholders based on country specific definitions. Where European data included a > 100 ha category, we included this in the 100-200 ha range, making our classification less precise in > 100 ha groupings, in comparison to < 100 ha. Future researchers may wish to aggregate all ‘large’ farms into a > 100 ha bin for their specific needs, but here we present the results maintaining the disaggregation for surveys that reported it.

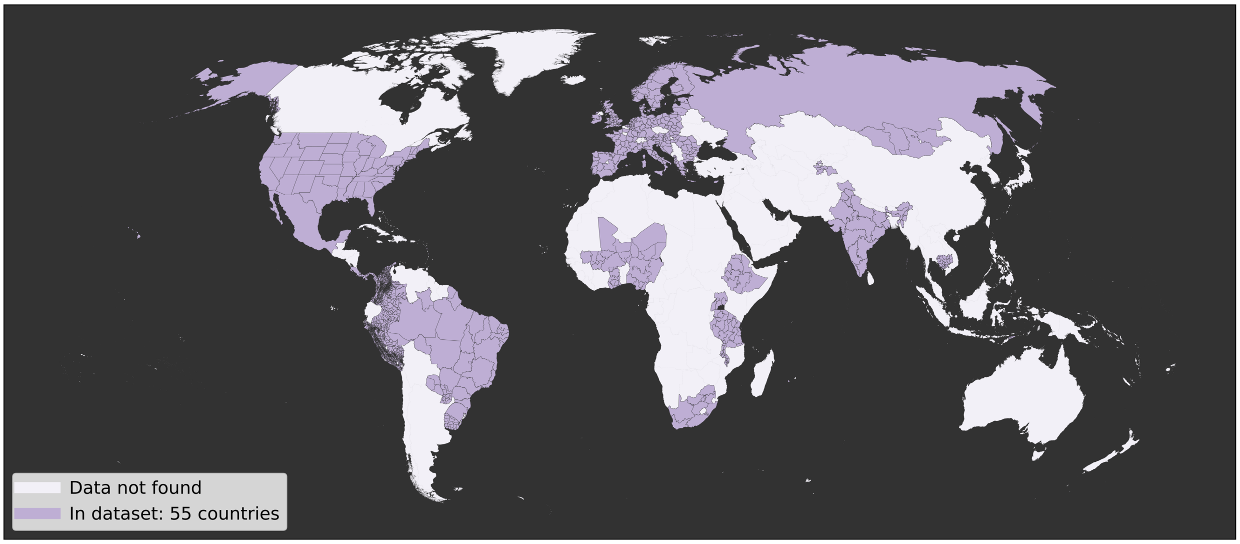

Figure 3: Spatial coverage and resolution of our data on crop production by farm size. Countries shaded purple had directly measured data on crop production or harvested area.

2.3.2 Crop allocation

Following data compilation, we converted all tonnes of production to their kilocalorie (kcal/capita/day) equivalents using FAOSTAT conversion values per crop per country per year. We then applied the percent of feed, food, processing, seed, waste, or 'other' based on FAOSTAT's food balance sheets per crop per country per year. For example, in many countries, maize can be used for human consumption, animal feed, a processed biofuel commodity, and seed, while some maize may be lost due to storage and transportation. FAOSTAT contains national totals for each of these types of crop allocation categories. We used these totals to calculate percentages per crop per country per year to allocate a certain portion of each crop's production towards food, feed, and the other crop allocation categories. While this approach does not account for the actual distribution of crop allocation by farm size, it is the most detailed information available and represents a proxy indicator based on what type and quantities of crops each farm size produces.

While certain FAOSTAT categories were straightforward to interpret and contained detailed definitions (e.g., 'feed' towards livestock and poultry and 'seed' set aside for sowing or planting), the processing category was ambiguous and required us to make assumptions. We followed Cassidy et al. and assumed that the processing category included oil crop production into oils for human consumption and for industrial use, as well as protein dense cakes for animal feed (43). The waste category encompassed any loss of a given commodity during storage and transportation; losses incurred before and during harvest were excluded, as were losses due to household consumption. The 'other' category encompassed any uses not already accounted for.

After allocating all crop production to type of production (e.g., feed, food, other, etc.) in kcal/capita/day we evaluated how the global quantity of each varied across farm size classes. We also provide cumulative distributions of our estimates to encompass a sliding scale of definitions of small-farms (e.g., farms under 2 ha, farms under 50 ha, etc.), as may be required by different researchers and regional policy makers who might define 'small' using different thresholds. In addition to comparing how the type of production varies across farm sizes, we also analyzed how the types of production are distributed within each farm size class.

To obtain global estimates for the proportions presented in this manuscript, we computed 95% confidence intervals using the accelerated bias-corrected percentile limits bootstrap method (BCa), with 1000 iterations. BCa is useful extension of the basic percentile bootstrap, that decreases coverage error by accounting for bias in sample parameters (i.e., when the sample parameter - computed from the 55 countries -- does not equal that of the average bootstrapped parameter, which in our case is our best estimate of global trends), and allowing the standard deviation of the bootstrap parameter to vary with the sample parameter (44). We chose to bootstrap all of the parameters of interest at the level of the country (n = 55), not the administrative unit (n = 3410), in an attempt to account for dependencies amongst administrative units in the same country and sampling campaign. Accuracy in uncertainty estimation for the global trends could be improved in future by adding to the number of countries in the dataset. While the BCa does not make any assumptions about the distribution of underlying random variable we use the natural log transform of production in our analysis for data visualization.

2.3.3 Crop species diversity and crop types

To estimate the relationship between crop diversity and farm size, we counted the proportion of unique number of species each farm size category produced within each administrative unit, and estimated the 95% CI's for each category using BCa. We note that different survey instruments have different crops included, and that farmer responses may not include the full diversity of crops that farmers actually produce. Thus, our estimates represent our current state of knowledge given empirical data, and are likely to be conservative. We present BCa estimates of crop diversity for each farm size using the administrative unit level to compare crop diversity distributions across farms within similar biogeographical landscapes (e.g., climate, soil, etc.). We also present BCa estimates of crop diversity while controlling for cumulative farm area, to give an indication of how diversity scales across the world in each farm size class. To do this we plotted cumulative numbers of unique species against cumulative area of administrative units for each farm size class, and estimated uncertainty for these curves by resampling the distribution of administrative units for each size class at random, with 1000 iterations (taking the 2.5^{th} and 97.5^{th} percentiles as lower and upper bounds, respectively).

To examine the variation in crop groups by farm size, we aggregated our crop species data into major commodity groups according to FAOSTAT definitions of cereals, fruit, oil crops, pulses, roots and tubers, tree nuts, vegetables, and other, and we estimated 95% CI's using BCa. Relying on the FAOSTAT classification has its limitations. For example, soy was classified as an oil crop, but it is also a pulse; therefore, this classification should be used as a guideline (see Appendix A for crop grouping details). In order to examine whether different farm sizes grew a different portfolio of crop groups, we used Sorensen's similarity index:

$$CC_{i,j} = \frac{2C_{i,j}}{S_i + S_j}$$

$C_{i,j}$ is the number of species two farm size classes have in common, S is the total number of species found in the given farm size class, and i and j are the two farm size classes being compared; a score of 1.0 would represent perfect overlap in the crop groups grown between the two farm size classes.

2.3.4 Macro-nutrient production

We converted production of each crop in our dataset to its macro-nutrient (i.e., carbohydrate, protein, or fats in grams/capita) equivalent using FAOSTAT food balance sheets, and conversion factors per crop per country for the year matching the farm size data survey year. Any temporal data gaps in FAOSTAT were linearly interpolated per crop and country. As with production, we analyzed how macro-nutrient production varied both across farm-size classes and within farm-size classes and computed 95% CI's using BCa at the country level to estimate global figures.

2.4 Results

2.4.1 Crop allocation

The smallest two farm size classes (0-1 ha and 1-2 ha) are the greatest contributors to global food production compared to all other classes. Farms less than 2 ha produce 28-31% of total crop production and 30-34% of the global food supply (by calories; Figure 4 A-H) as extrapolated from the 55 countries in our dataset. Their contribution is slightly higher than their areal coverage of 24% of gross harvested area, suggesting small farmers have greater cropping intensity or higher yields than larger farms.

We found smallholders (farms < 2 ha) also allocate the largest percentage (55-59%) of their crop production to food compared to all other farm size classes (Figure 4). Generally, larger farms devote more of their production towards feed and processing. Farms between 200-500 ha have the largest allocation of their production to feed (16-29%) compared to farms < 2 ha who allocate 12-16% to feed. Farms > 1000 ha allocated 12-32% of their production to processing.

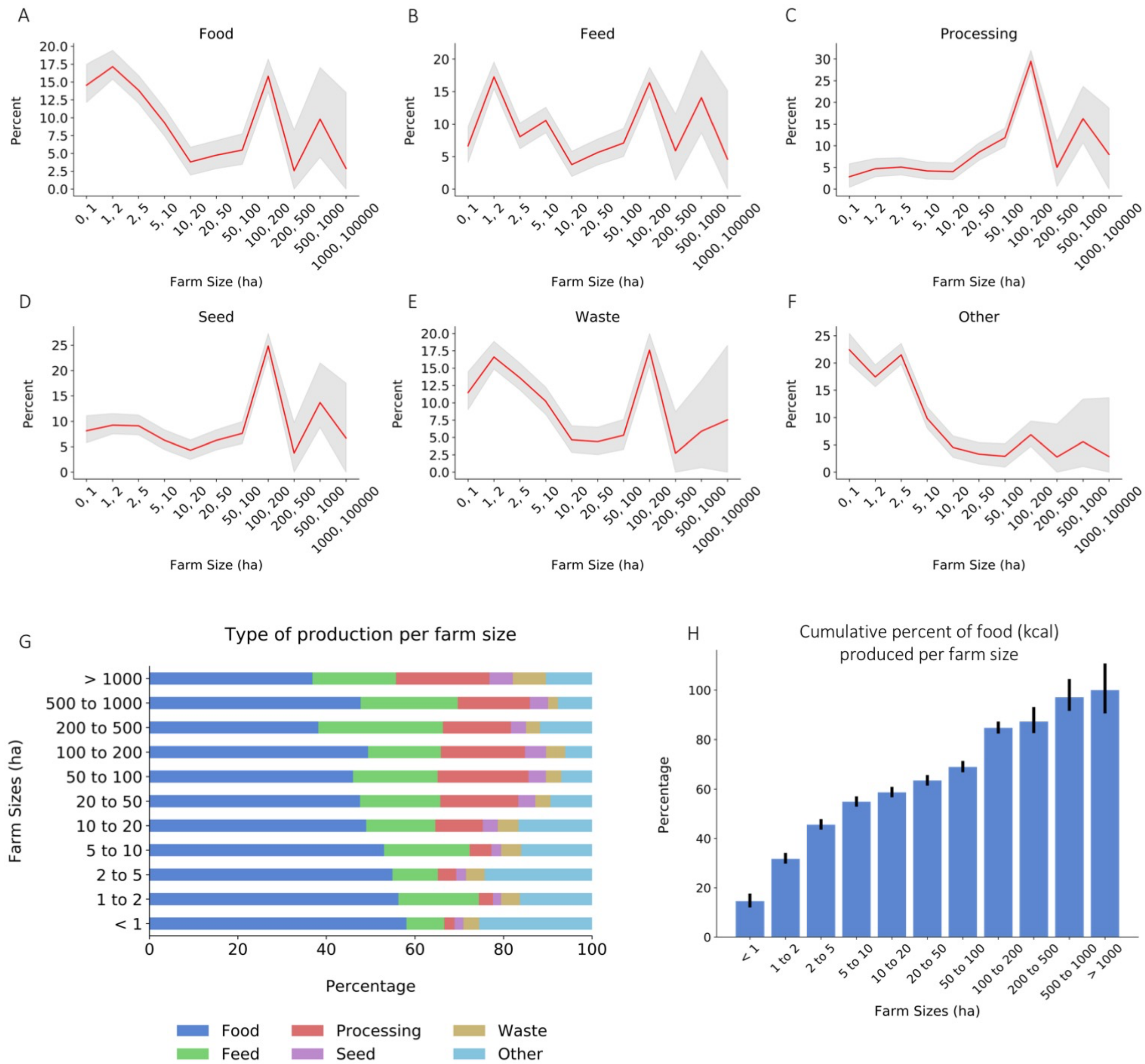

Figure 4: A-F) Distribution of total global crop production (in kcal equivalents) across farm size groups different uses (e.g., food, feed, other, etc.). Grey shows bootstrapped 95% confidence intervals and red indicates the average. G) Allocation of use of production within each farm size class. H) Cumulative percent of global food production by farm size group with 95% confidence intervals. See Table S3 for underlying data.

Farms < 2 ha contribute the most 28.1% (26-30%) to total food waste (on-farm and post-harvest loss) (Figure 4E); however, this is mainly driven by this farm size group's large contribution to the total crop production. In our dataset, only 4% (2.3-6.1%) of smallholder production is wasted, compared to farms > 1000 ha that have the greatest amount of within farm size class waste at 7.5% (0.0-18.5%). However, the large uncertainty indicates both that there is substantial variation within large farms, and low confidence in the trend between farm size and waste holds at the global level. All farm sizes have fairly consistent allocations towards seed (means ranged from 2-5% with overlapping 95% CI's), while there is a trend that smaller farms allocate more to the 'other' category.

2.4.2 Crop species diversity and crop types

We found that species richness declined with increasing farm size (Figure 5). Diversity also scaled differently with area within different farm size classes, with greater turnover in unique species in small farms than in land allocated to larger farms.

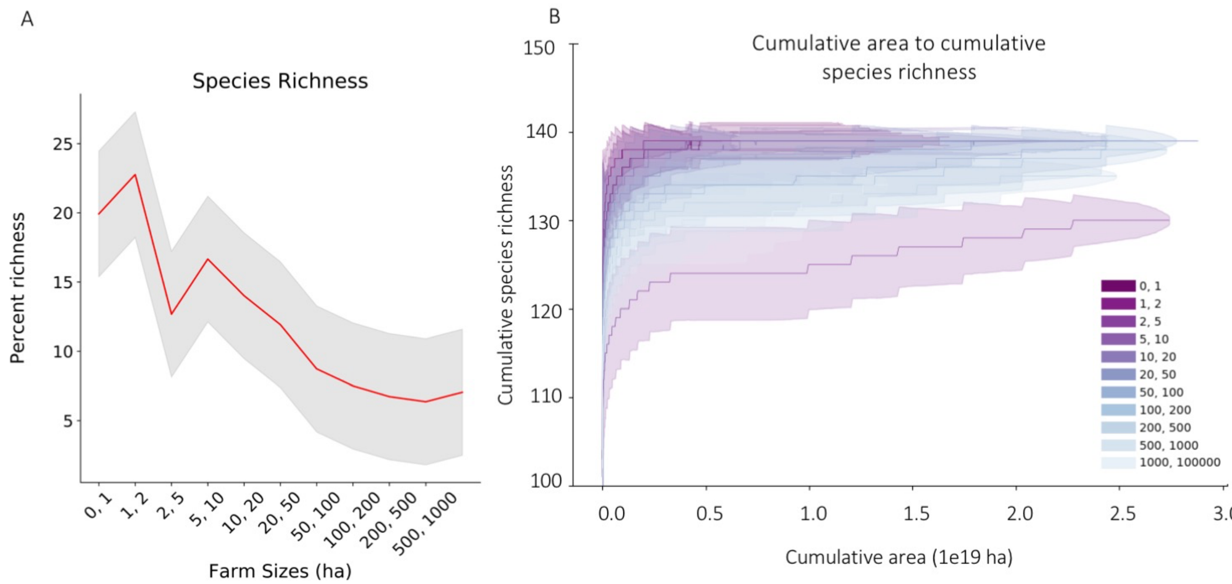

Figure 5: A) Distribution of total species richness across farm size classes. Gray represents bootstrapped 95% confidence intervals; red is the bootstrapped average. The area of each administrative unit polygon weighted the data. See Table S4 for underlying data. B) Cumulative area to cumulative species richness curves. 1000 iterations generated the cumulative distributions between species richness and farm size. The starting point for cumulative distributions were randomly chosen each iteration. The lighter the colors, the larger the farm size classes.

Between farm size dissimilarity in species shows that larger farms, while harboring less diversity, and lower turnover in crop diversity across space, show greater specialization in certain crop groups than other farm sizes. Farms < 5 ha grow similar crops as each other (Sorensen's coefficient of 0.94), and farms > 100 ha have a perfect overlap in crops grown (Sorensen's coefficient of 1.0; Figure 6). But farms greater than 20 ha grow a different array of crops compared to farms smaller than 20 ha (Sorensen's coefficient of 0.4-0.67) and farms greater than 100 ha have the lowest overlap with other farm size classes.

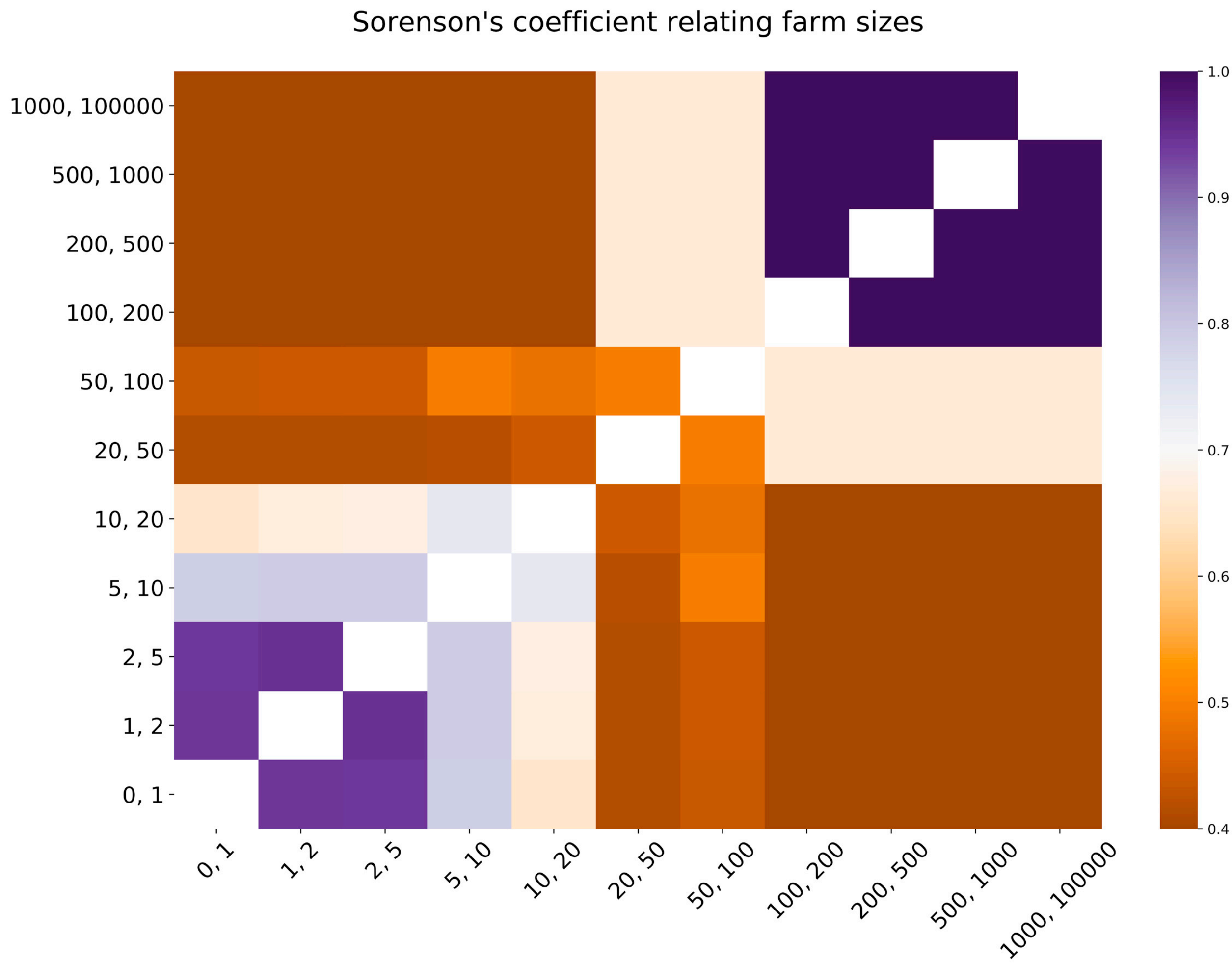

Figure 6: Heat-map of Sorensen's coefficient between each farm size class pair. Purple indicates a greater similarity of crops grown between pairs of farm size classes, while brown indicates greater dissimilarity between crops grown.

The crop portfolio of each farm size class shows that smaller farms (< 2 ha) produce a greater share of the world's fruits, pulses, and roots and tubers, while medium sized farms produce more vegetables and nuts, and large farms produce more oil crops and 'other' (Figure 7). While all farm sizes contribute a large proportion to cereals, smaller farms devote a greater percentage of their overall production to cereals compared to other farm size classes.

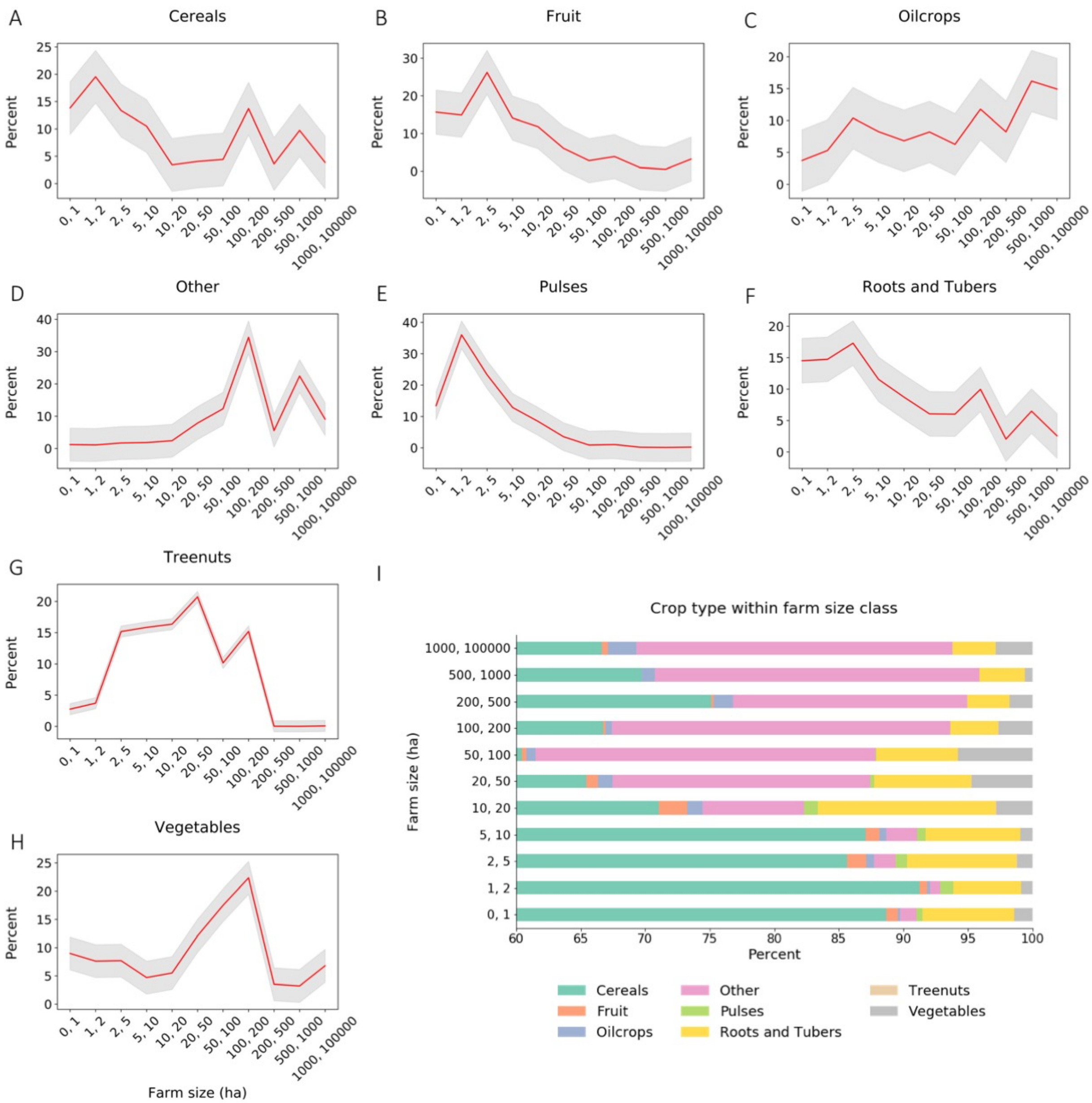

Figure 7: A-H) Distribution of production by crop type across farm size classes. Grey shows bootstrapped 95% confidence intervals and red is the average. I) Crop type portfolio within each farm size class; x-axes ranges from 60-100% since cereals were the only group across all farm size classes < ~60%. See Table S5 for underlying data.

2.4.3 Macro-nutrient production

The trends in macro-nutrient (carbohydrates, proteins, and fats) production across farm sizes follows that of the food production. Yet, of their own production, smaller farms produce a slightly higher percentage of carbohydrates (~0.08% more than the largest farm size class) while larger farms grow a slightly higher percentage of proteins (~0.05% more than the smallest farm size class). But these differences are minute, and considering the uncertainty estimates, there are no significant differences in the percentage of macro-nutrients produced within each farm size class (Figure 8).

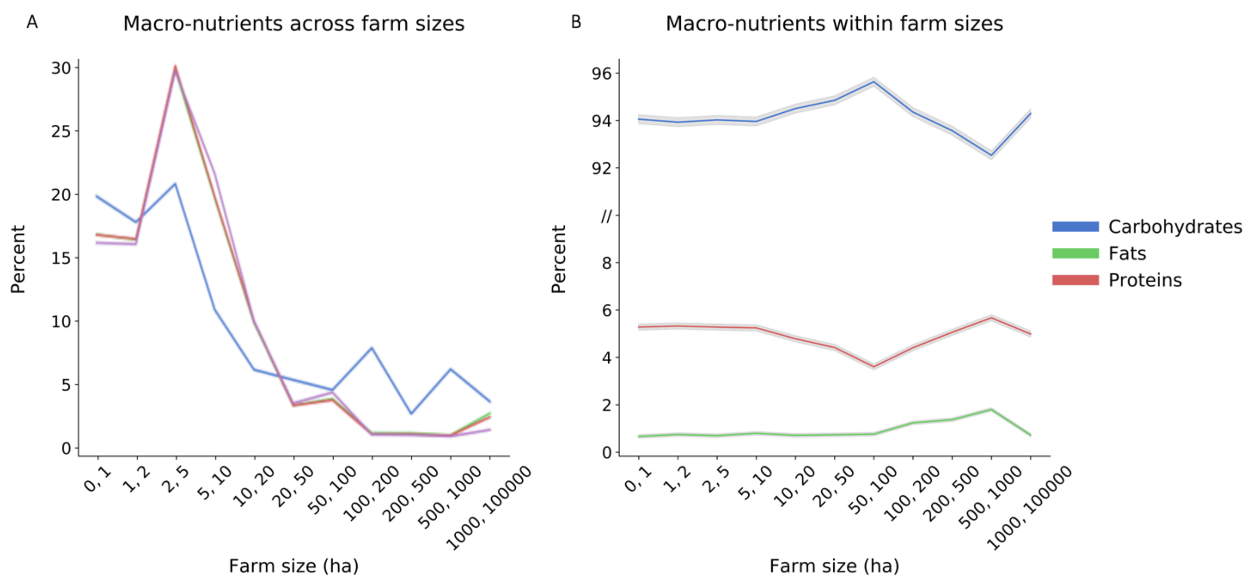

Figure 8: A) Percentage of macro-nutrient production across farm size classes with 95% confidence intervals. B) Percentage macro-nutrient production within each farm size class with 95% confidence intervals. See Table S6 for underlying data.

2.5 Discussion

2.5.1 Comparison to previous studies

Our dataset is the first global sample of direct crop-specific measurements of production or area by farm size. We found that farms < 2 ha produce 28-31% of total crop production and 30-34% of the food supply on 24% of gross agricultural land when using our directly measured farm size dataset. While our dataset covers 55 countries, with distinct data gaps in smallholder dominant Southeast and East Asia, our findings are in line with Samberg and Herrero's global estimates. This suggests that these three studies, using different methodologies, agree that the previous estimate of smallholders producing 70-80% of global food production needs to be revised.

While our results are similar to the previous two modeling studies that estimated global smallholder production, there are several key differences. Our results offer more refined estimates using direct measurements of production by farm size instead of relying on modeling, includes a larger range of crop species than previously assessed, and our accompanying open-access dataset allows individual countries to have a reliable SDG baseline for how much of their food production is grown by smallholders (according to their own regional definitions of farm size). Samberg reported that farms < 5 ha produced 55% of global food calories, which is slightly larger than our equivalent estimate of 44-48% (Table 2, Samberg A.). To arrive at this estimate, they divided the total calories produced in each farm size category in their 83-country sample by total global calories produced by all countries. Their estimate could be considered a global estimate if one assumes that their sample of smallholder dominant regions account for most of the world's small farms (and their purposeful sample might suggest that interpretation). An alternative interpretation (which is similar to ours) is that the 83 countries in their sample are globally representative; in that case one would divide the calories produced by each farm size class by the total calories produced in those 83 countries. By this estimate, farms < 5 ha produced 76% of global food calories (Table 2, Samberg B.). The differences between Samberg and our dataset may be due to the countries and crops sampled and that Samberg relying on modeled results instead of direct measurements. Samberg used 41 crop species (while we included 154), and they use mean agricultural area instead of farm size distributions to understand crop production in smallholder dominant areas rather than crop production by farm size. Both Samberg and our study relied on household sample surveys to varying degrees (Samberg relied primarily on household surveys, while our dataset relied on them for 22.5% of total crop production). Household sample surveys systematically do not sample non-family farms and hence may be presumed to over-represent smaller farms when compared to agricultural censuses that survey all farm types. However, in our Appendix A we show that using household surveys to estimate national production is not significantly different than using FAOSTAT's national production estimates.

Table 2: Comparison between global estimates for the percentage of food smallholders produce. For Samberg, A. uses estimates compared to global total food production, while B. compares estimates to total food production within their 83 sampled countries.

	< 2 ha	**< 5 ha**	**< 50 ha**	**Methodological Distinctions**
Our study	30-34%	44-48%	62-66%	Direct measurements 154 Crops 55 Countries
Herrero	20%	-	51-77%	Modeled estimates 41 Crops; 7 Livestock; 14 Aquatic Species Near global coverage
Samberg A.[1]	37%	55%	-	Modeled estimates 41 Crops 83 Countries Mean Agricultural Area as farm size proxy
Samberg B.	52%	76%	92%	

Our estimates were also close to Herrero who reported global estimates for farms < 50 ha. We found that farms < 50 ha produce 62-66% of the world's food, which is within Herrero's range of 51-71%. Herrero found that farms < 2 ha produce ~20% of food. Our two studies capture different aspects of the global food system. Herrero incorporated livestock and fisheries, which are important source of nutrients and income for smallholders, while we only focus on crop production; our focus was due to data constraints and definitional mismatches between using farm size versus herd size, fishing area, or common pasture land. There are also crop species differences between the datasets, where Herrero used 41 crop species while we used 154. One key analytical difference is that Herrero's modeled results used field size as a proxy for farm size instead of actual

1 Note that we do not provide < 50 ha estimates for Samberg A because we cannot support the assumption that there are no farms < 50 ha outside of the 83 countries sampled by Samberg.

reported farm size; we used field size as a proxy for farm size for only 4.8% of our data and direct measurements of production by farm size instead of modeled estimates. We found that using field size as a farm size proxy measure may slightly over estimate small-farms' production since it does not account for non-field elements of a farm (see Appendix A). Additionally, Herrero disaggregated production to the pixel level based on field size, while Samberg disaggregated pixel-level production based on mean agricultural areas. Essentially, both methods assume a constant yield for each farm size class since they cannot directly link crop production with farm size. There is a widely observed inverse relationship between yield and farm size (IR), where smaller farms have higher yields. For 66.7% of our dataset we also needed to use constant yields since direct data on production by farm size was not always available (we did have harvest area per crop by farm size, and minimally used data on planted area, cropped area, plotted area). Our dataset allowed us to test for the bias introduced by constant yield methods, and provides the relationship which researchers may use to correct for it. We found a small effect size that using constant yields slightly underestimates small-farms' production (Table S2). Hence, our numbers, Herrero's and Samberg's may all slightly underestimate smallholders' crop production owing to this assumption.

2.5.2 Crop allocation

Our new findings on crop allocation across different farm sizes has important implications for food access and availability, as well as farmer livelihoods, since food, feed, processing, and seed market prices may differ from one another. We found nearly 60% of smallholder production is allocated to food. A smaller percentage is allocated towards feed (12-16%), which was surprising since smallholders often engage in mixed crop-animal farming systems (45); this finding may be explained by the fact that smallholders likely rely more on rearing animals that graze on pasture compared to largeholders.

Our results counter common thinking about smallholders' post-harvest loss, where improving cold storage and road infrastructure is a common development intervention to improve smallholder income by reducing wastage. Our dataset suggests that only a small percentage of smallholders' production is wasted. However, one reason for the low amount of smallholder waste in our results may be due to food allocated to the 'other' category. From our data, 19-23% of smallholder production went towards 'other' uses. This may be indicative of the need for smaller farms to make use of all grown material in integrated farming systems (e.g., using rice stocks as a cover crop to promote soil health). Smallholders' large allocation towards 'other' may be indicative that waste reduction practices are common since smallholders are often resource poor and would achieve higher relative benefit compared to largeholders to find a use for wasted crops.

While, interventions aimed to reduce smallholder post-harvest loss are still needed in many locales, there is also a need for agricultural researchers to identify why larger farms are wasting crops, because this group showed the greatest proportion of waste in any category (although this was country dependent, as shown by the wide bootstrapped confidence intervals). An estimated 1/4 of global food production from croplands is wasted from farm to market (46). The waste data we used takes into consideration the quantities lost in the transformation of crop to processed goods. Hence, one possible reason for the increased wastage of larger farms' is that large farms on a whole engage in more crop production allocated for processing. Since FAOSTAT's definition of waste also encompassed waste incurred from poor distribution and storage, it was surprising that smaller farms did not have larger proportions of their crop wasted when the majority of smallholders are in countries with hot and humid climates and poorer storage infrastructure (47). Future studies should disaggregate the types of crop production waste each farm size contributes to and local dependencies for these relationships.

2.5.3 Crop species diversity, crop types, and macro-nutrient production

Our data suggests a negative relationship between farm size and crop species richness. This adds significantly to the evidence on the literature's mixed finding on this relationship (40, 41, 48), as our study contains a wider range of farm sizes and more crop species than ever compared in previous studies. Due to the heterogeneity in data sources used to construct our dataset, there were not always a wide list of crop species included in each national survey, which may indicate a larger portion of primary crops were documented compared to local species. This limitation indicates that our findings are conservative and suggest that smaller farms, which are associated with producing many non-primary crops (49, 50), may have even higher degrees of crop diversity than we found.

There are several food access, nutrition, and climate resilience implications of higher crop diversity in smallholder systems. Since smallholders may be tied to subsistence-surplus production models and constrained to highly localized rural markets, their food access is often more reliant on their local communities' crop production compared to large farms (51); hence, in these local markets, farms' crop production needs to be more diverse than better integrated markets that can rely on imported crop diversity. The differences in types of crops produced by different farm sizes, and macro-nutrient contents that follow overall food production trends, supports the differences in macronutrient production across farm sizes, as found in Herrero. However, there are discrepancies. While smallholders produce a large amount of the world's protein rich pulses, we did not find that they produced a greater relative percentage of proteins than larger farms (i.e., all farm sizes allocated a similar percentage of their production to proteins). This suggests potential benefits from the promotion of mixed animal-crop systems for smallholders to access protein of which they are often deficient.

Our results suggest a nuanced view of the benefits of landscapes harboring different farm sizes, beyond the basic relationship between farm size and crop species richness. More diversified farming landscapes may need to include smaller farms because they collectively grow a higher diversity of crops than large farms, but also include larger farms because of their unique crop composition. Each farm size produces a greater quantity of certain types of crops than other farm sizes: smaller farms produce more fruits, pulses, and roots and tubers, while medium sized farms produce more treenuts and vegetables, and larger farms produce more oil crops. Promoting a diversity of farm sizes may encourage a greater diversity of crop types at the landscape level that can better provide more balanced diets and non-food needs, while potentially mitigating climate risks to the food system as a whole since different species are not affected by temperature and precipitation changes the same.

2.6 Conclusion

This study attempted to provide a global baseline for international policy measures aimed to support smallholder agriculture. These include a need for improved monitoring of SDG Goal 2.3, which aims to double food production of smallholders and increase nutrient availability; yet, Goal

2.3's monitoring framework does not use crop production by farm size as a national indicator (10). Our findings suggest that previous estimates of the percentage of food produced by smallholders were either overinflated by public-sector opinions or still needed directly measured data to assess quality (Herrero; Samberg), and that a nutrient diverse farming landscape would include a diversity of farm sizes, since each farm size produces a unique crop portfolio.

Critically, while our dataset is the first to use directly measured crop specific data on production or area by farm size, we were only able to find 55 countries with the necessary data to do this analysis. To monitor SDG Goal 2.3, there needs to be increased effort to build on datasets like ours through leveraging stakeholder networks. Ongoing efforts to use and add to our dataset will enable continuous food system monitoring over time with more geographic precision. We urge researchers and food system advocates towards data-driven policy monitoring to accurately assess the scale and progress of policy interventions.

Chapter 3: Smaller farms are higher yielding and more biodiverse than larger farms: A systematic review and meta-analysis

3.1 Abstract

The scale of agricultural production is rapidly changing. In general, farms are getting larger and more industrialized in high-income countries and smaller and more fragmented in low-income countries. There has been limited empirical synthesis of the implications of changing farm size for outcomes related to food security, economic development, and environmental sustainability. To identify and assess the multiple trade-offs and context-specificities influencing these relationships, we present the first systematic review and meta-analysis on how differences in farm size affect crop production, farmer livelihoods, and the environment. We analyzed 118 empirical studies (318 observations) from 52 countries, and show that smaller farms, on average, have higher yields (5% decrease in yields per 1 ha increase in farm size), promote non-crop biodiversity at the farm and landscape scales (77% of studies find smaller farms have more non-crop biodiversity than larger farms), and account for greater crop diversity than larger farms (farms under 2 ha accounted for ~40% of crop species richness in a given landscape). We find no strong relationships between farm size and resource-use efficiency, greenhouse gas (GHG) emissions, or profit per ha. Our findings highlight the importance of farm size in mediating environmental and social outcomes relevant to sustainable development and identifies a series of research priorities related to the ongoing global transition in farm sizes.

3.2 Introduction

The scale of agricultural production is rapidly changing. In general, farms are becoming larger and more industrialized in high-income countries, and smaller and more fragmented in low-income countries (1). The primary factors mediating changes in farm size are economic development, land consolidation and redistribution policies, traditional land inheritance systems, and Large-Scale Land Acquisitions (LSLA) (8, 9). Yet, across contexts the cumulative impacts of farm size transitions remain largely underspecified for both farmers and wider society.

The majority of the world's farms are still small -- of the 570 million farms in the world, 84% are less than two hectares (ha) in size and constitute 12-25% of farmland area (1, 29). Many scholars claim that smaller farms are more productive, resource-use efficient, and environmentally friendly than larger farms (2–4), and yet farmers operating smaller farms (smallholders, hereafter) are facing growing pressure on their livelihoods from low prices in global markets and climate-change induced production losses (5). As a result, small farms are a central focus of international organizations promoting sustainable development agendas around improving livelihoods, food security, and environmental health. For example, the United Nations (UN) Food and Agricultural Organization (FAO) actively promotes small and family farms (e.g., the FAO's "Year of the Family Farm" in 2014). The 2015 Sustainable Development Goals (SDG) seek to support smallholders by increasing their productivity, incomes, and access to land (SDG 2.3), while the 2015 UN Conference of Parties on Climate Change (COP21) agreement resulted in a wide array of commitments to bolster smallholders' adaptive capacity (52). At the same time, consumers are increasingly concerned about health, farmer livelihoods, and the decline of diverse and traditional foods, leading to an increased willingness-to-pay for products with organic and/or local labels, which are often associated with smaller farms (13–17).

But despite the international policy activity around farm size, little research to date has systematically examined how farm size relates to both environmental and socio-economic outcomes. On the one hand, isolated case studies suggest that smaller farms can be more resource-use efficient, productive, and biodiverse than larger farms (53–55). While the observation that smaller farms exhibit higher yields than larger farms was first studied in the 1960s (56, 57), there has been little systematic work to understand how the relationships between farm size and outcome variables vary across levels of economic development, type of production system, and farm size ranges. A robust and multi-dimensional perspective is needed to better understand the economic, social, and environmental impacts of changes in farm size globally.

Here, we synthesize the relationships between farm size and select sustainability outcomes across a range of geographies, leveraging the past 50 years of empirical evidence that directly assessed crop production, environmental performance, and economic outcomes as they relate to farm size. Our systematic assessment of multidimensional outcomes of farm size builds upon past reviews focused on single outcomes (e.g. yield, economic performance, or biodiversity metrics for specific species) (23, 58–60); non-systematic reviews (58, 59); studies based on indirect measurements of farm size and the outcome variables of interest (18, 47); and studies with specific regional foci (58, 60). We build on these previous efforts by including studies across diverse regions that directly assessed farm size and the outcome variable(s). We assess six relationships between farm size and: 1) yields (value of crop output per area (value/ha), and total crop production per area (kg/ha)), 2) non-crop biodiversity (field and landscape levels), 3) crop diversity (species and varietal levels), 4) resource-use efficiency, 5) greenhouse gas (GHG) emissions per unit output, and 6) profit per unit area. For yields, resource-use efficiency, and GHG emissions (for which we had quantitative data), we use meta-regressions to compute pooled estimates for these relationships and examined each variable in relation to location, crop, and farm management contexts. We bound this review to spatial definitions of size (i.e., farm size in terms of operated area) as it is a central dimension of the farm size debate.

We situate these results across diverse development contexts to identify if there are lessons from one or more context(s) that may provide policy-relevant insights for others. Our findings show that smaller farms have higher yields, promote non-crop biodiversity at the field and landscape scales, and account for greater crop diversity than larger farms. We find no strong relationships between farm size and resource-use efficiency, GHG emissions, or profit. Our findings also highlight a series of important literature gaps, with implications for future research aimed at balancing equitable economic development with environmental impacts.

3.3 Results and Discussion

3.3.1 Literature search

A systematic review and meta-analysis was conducted using the PRISMA guidelines (61) (see Figure S12 for inclusions/omissions and Table S7 for Boolean search terms). The search resulted in 218 observations (from 111 studies) for a "vote-count" analysis, designed to estimate the probability of studies finding a negative, null, or positive relationship between farm size and the outcome variable (Table S8 shows summary statistics). Studies often contained multiple observations and levels of detail if they separately measured different outcome variables, crops, or locations. Of these, 70 observations (in 45 studies) had information to allow for a meta-analysis of regression coefficients to estimate the relationship between farm size and socio-environmental outcomes. To augment the sparse crop diversity and GHG emission literature on farm sizes, we used results from Ricciardi et al (2018) and Clark and Tilman's (2017) dataset, respectively; the latter consisted of 100 observations from 11 studies we used in our analysis, for an overall total of 318 observations. As part of our systematic review we extracted information from the broader literature on causal mechanisms behind the main trends as well as factors that caused deviations from the main trends (see Table 3). Below we present both the findings from our systematic review and meta-analysis.

Table 3: Main results and mechanisms.

Variable	Result	Mechanisms Benefiting Small Farms	Mechanisms Benefiting Large Farms
Yield	Smaller farms	• Reliance on family labor (e.g., Figure 10).	• Precision agriculture mechanization enables higher yields with less labor, but is only cost effective on larger fields (62).
Biodiversity (non-crop)	Smaller farms	• Smaller fields have more edges that provide habitat (63, 64). • Independently managed smaller fields create a more heterogeneous landscape (65).	• The link between field and farm size is relatively understudied; large farms with small fields may also benefit biodiversity, but was untested in the reviewed literature.
Crop diversity	Smaller farms	• Subsistence farmers plant a greater diversity of traditional crops to meet nutritional needs (49). • Small farms are incentivised to cultivate landraces when there are niche markets for traditional crops (50).	• Varietal diversity requires a minimum amount of space to prevent genetic erosion for wind-pollinated crops (66, 67). • Diversified crops can reduce long-term risk at the expense of short-term profit (48, 68).
Resource-use efficiency & GHG Emissions	No relationship	• In contexts where off-farm labor opportunities were greater, there was less available on-farm family labor and, in turn, greater technical efficiency (69). • Smaller farms may use less input intensive production methods, but was untested in the reviewed literature.	• Agriculture mechanization can enable higher yields with less labor and more efficient input use, but is often only cost effective on larger fields (62). • Increased access to information from extension and advisory services was associated with highly efficient farms, but was more accessible to larger farms (69–72).
Profit	No relationship	• Specialty markets for traditional foods offers higher prices (50). • Smallholders' credit access can increase access to inputs and markets (73).	• Better market access for larger farms (76, 104). • Recovering fixed costs require a minimum scale (74, 75). • Land-based subsidies (76).

3.3.2 Yield

The inverse relationship (IR) between farm size and yields -- a trend describing how production per unit area declines with increasing farm size -- has been widely studied since the 1960s (56, 57, 77–79). Microeconomic theory suggests that optimal sizes exist for specific production processes (80), which has led to the debate on whether there is an optimal farm size for specific crops in different political and economic contexts, and if policies should thus redistribute land to increase regional productivity (9, 77, 81, 82).

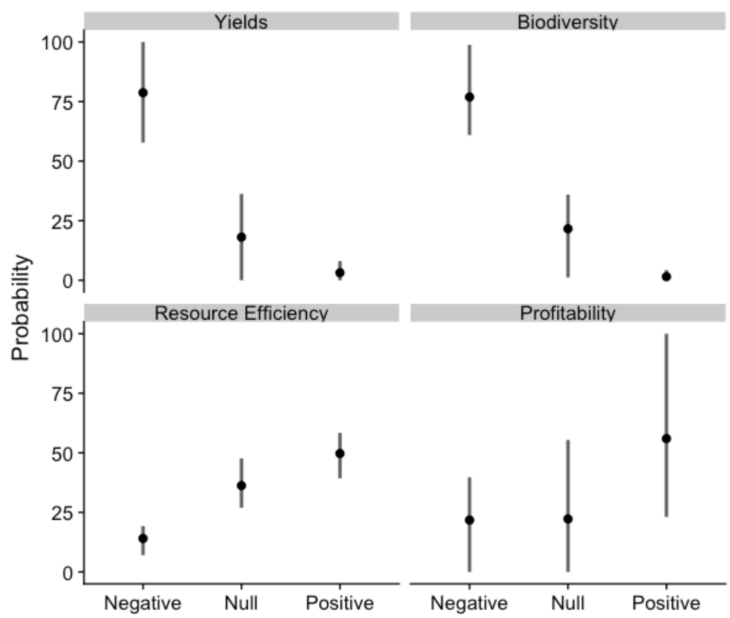

Figure 9: Plots A-D show the probability of studies finding negative, null, or positive relationships between farm size and the outcome variable as per the vote-count findings. The average and 95% confidence intervals are given. (See Table S9 for underlying data.)

Our analysis found that the IR holds across a wide array of country contexts and crop types. We estimated that 79% of studies (95% CI = 58-100%) find smaller farms to have higher yields (in both weight/ha and value/ha terms) than larger farms (Figure 9). We also examined pooled effect sizes to estimate the magnitude of change in yield per one hectare change in farm size. The pooled estimates were computed from the extracted and standardized regression slopes (where studies regressed farm size onto yield). We found that yields typically decreased by 5% for each hectare increase in farm size (i.e., -5% mean effect; 95% CI: -9 to -1%; Figure 10). These results were consistent across measurements defining yield as weight/ha or output value/ha. While confidence intervals around these findings are large, and the distribution of effects includes non-consistent cases, these findings show that, on average, the available evidence supports the hypothesized IR.

In 1964, Sen offered three explanations for the IR -- differences in farming practices, labor markets, and/or land heterogeneity -- that continue to be tested for different crops across a variety of institutional, labor, and environmental contexts (83). We explored if our results were moderated by these three common explanations in the literature through sensitivity analysis and based on the available evidence.

Additionally, there were no consistent differences between studies when they controlled for farming practices (Figure 10). While not included in our review due to our inclusion criteria, precision agriculture mechanization has been found to enable higher yields with less labor, but is only cost effective on larger fields (62). Our findings, along with the limited studies we found on yields and farm size in higher-income countries, suggests that larger farms may disproportionately be able to invest in mechanization to overcome the IR as optimal farm structures have been found to change with economic development (59).

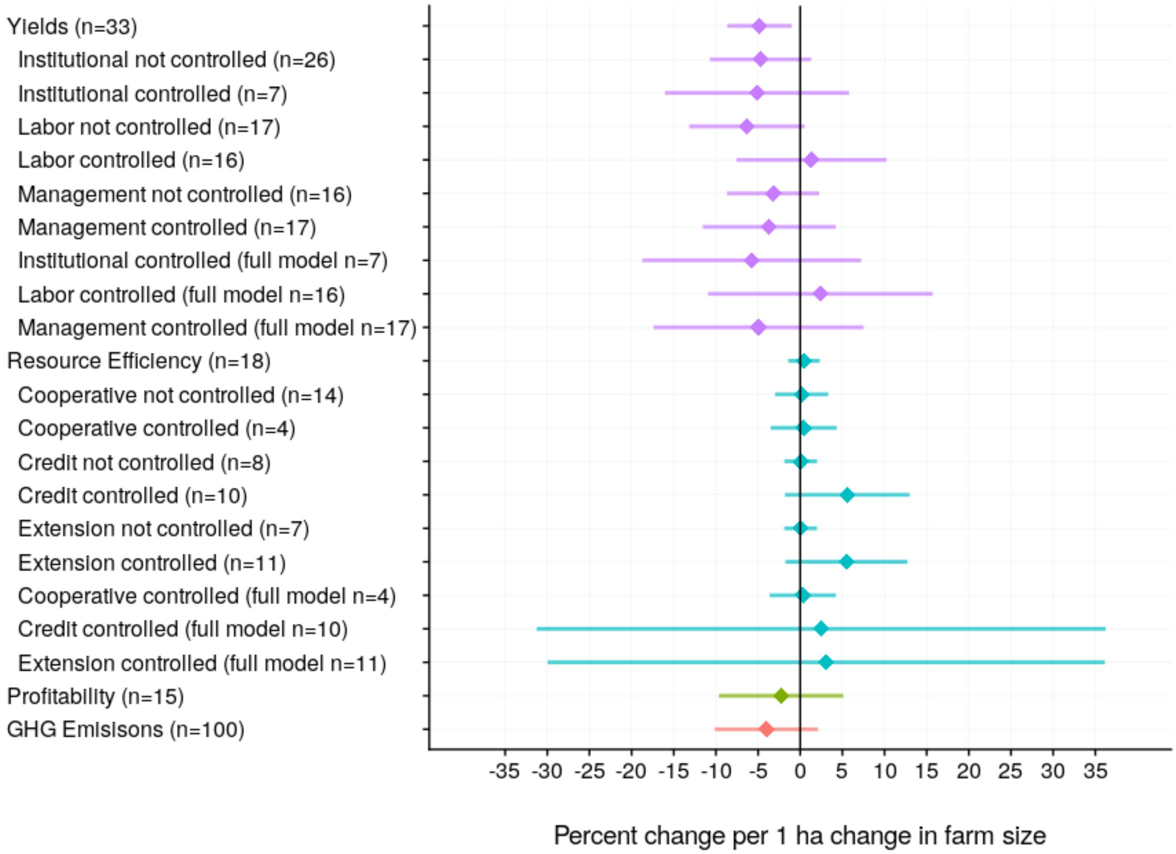

Figure 10: Pooled effect size per variable as derived from the random effect meta-regressions. The vertical black line indicates the 1:1 response ratio where for a 1 ha change in farm size, there is no change in the variable. If a response ratio is < 0, then smaller farms have a higher effect (e.g., smaller farms have higher yields), and if it is > 0, then larger farms have a higher effect. The number of observations (n) and 95% confidence intervals are given per variable. For yield, sensitivity analyses fit separate models to test if the effect was consistent for studies that controlled for common explanations for the inverse farm size to yield relationship: institutional characteristics, farm management, and family labor. For resource-use efficiency, separate models were fit to test if the effect was moderated by common development interventions to improve smallholder resource-use efficiency: extension access, farmer cooperatives/groups, and credit access. For yields and resource efficiency, full models indicate that all variables listed were controlled. Profit and GHG emissions had no additional models. (See Table S10 for underlying data.)

First, in exploring the impacts of institutional contexts, we found that the inverse relationship between farm size and yield held across studies controlling for credit markets, access to extension services, and involvement in farmer cooperatives.

Second, we found that for studies that controlled for types of labor (i.e., general labor market imperfections, family labor, and household size), the negative relationship between farm size and yield no longer existed (Figure 10). This may be explained by the fact that the presence of a large amount of unpaid family labor, or limited off-farm labor opportunities -- usually due to other market failures (e.g., unequal access to credit, land, and/or insurance) -- would lead to a high density of laborers on smaller farms. Controlling for the effect of laborers "boosting yields" on smaller farms leads to no differences in yield between farm sizes, which indicates labor is a key moderating factor to the IR.

Third, we found two rigorous studies on the relationship between land heterogeneity, yield, and farm size that collected soil samples in Madagascar rice paddies (55) and on Thai sugarcane farms (84); both studies found that soil differences did not affect the IR. Elevation and slope of fields have also been examined in India and in sub-Saharan African maize and grain farms, with the IR still holding (85, 86).

Lastly, it has been postulated that the IR is an artifact of measurement error, where farmers mis-report farm size and/or yield, but these hypotheses have had limited empirical work (87, 88). While the IR debate has measured productivity in terms of yield and various economic and production efficiency metrics (e.g., allocative, technical, and total economic efficiency, returns to scale, and profit), we discuss profit and resource-use efficiency separately below.

3.3.3 Non-crop biodiversity

Both smaller farms and landscapes dominated by smaller farms had greater non-crop biodiversity (across diversity metrics and species) than larger farms or landscapes dominated by larger farms. We estimated that 77% of studies (95% CI = 61-99%) find smaller farms to have greater biodiversity at the farm and landscape levels compared to larger farms. Since there were only 12 observations (4 studies) that measured farm size directly, we included and controlled for 75 observations (30 studies) that measured the relationship between biodiversity and field size, as field size has been shown to correlate with farm size (89, 90). While a key limitation to the generalizability of these results are that most studies were from high-income countries (Figure S13), the literature's key mechanisms may provide insight for other economic contexts.

Fields' "edge effects" are the key reported mechanism behind smaller farms having greater non-crop biodiversity. This is because field margins have been shown to contribute to non-crop biodiversity, and smaller fields have a higher field margin-to-field area ratio. Increased field perimeters (e.g., using grass buffer strips and hedgerows) lead to larger available breeding habitats for arthropods (91, 92), provide refuge for arthropods and smaller species to (re)colonize after escaping recently disturbed fields (63, 64, 93), increase the number of pollinators and beneficial predators within fields (64, 92, 94), and act as conservation corridors for arthropods and small mammals (63, 95, 96). While there is limited research on whether and how specific management practices influence the effect of farm size on biodiversity (23), there is evidence that farm size can have a larger effect than management on biodiversity. For example, Belfrage et al. found smaller farms to have more butterflies, birds, and non-crop herbaceous plants than large farms, irrespective of conventional or organic management (94).

At the landscape level, small farm-dominated landscapes also have greater biodiversity due to the buffers around their field edges providing wildlife corridors; diverse land cover types such as forest and wetland; fields of different crops; or fields in different stages of production (65, 97, 98). However, our review also found a few empirical cases where small farm and field sizes in a landscape still resulted in low overall landscape diversity. For example, Hansen and Libecap found that during the U.S. Dust Bowl, soil erosion rates were high at least partially due to the mosaic of small farms and fields which--although independently managed--had low collective crop diversity and land-cover types, resulting in a highly homogenized landscape (99). While the Dust Bowl is a unique example with diverse drivers, it reveals that simply having small farms does not inherently promote a biodiverse landscape; rather a diversity of management practices (including cropping patterns), fields, and biophysical contexts promote biodiversity at the landscape scale.

Not all species benefit from smaller farms and increased landscape heterogeneity. Birds have differing preferences for field size and landscape heterogeneity (100), and species-to-landscape heterogeneity relationships are not always unidirectional. For example, grey partridges in Poland use larger fields for nesting, but insect species that partridge chicks prefer as food are more abundant in smaller fields (101), suggesting that, at least in some contexts, there is a need for a diversity of farm sizes in the overall agricultural landscape to support biodiversity outcomes. Crop type and the landscape surrounding a field also had effects on bird populations, where certain migratory birds favored larger cereal and maize fields with close proximity to forested areas to access food sources, while other species preferred a small farm-dominated landscape (102).

3.3.4 Crop diversity

Many studies have explored *in situ* crop diversity among smallholder farmers (103–106), yet few directly measured the relationship between farm size and crop diversity. We only found eight observations (from eight studies, across six locations) that met our study inclusion criteria for crop diversity. Three of these studies found a negative relationship between farm size and crop diversity, while four found a positive relationship. We supplemented these vote counts, by instead relying on a more in-depth quantitative analysis from a recent publication where we addressed the relationship between crop diversity and farm size across 55 countries and 154 crops using a newly harmonized dataset of farmer surveys and agricultural censuses (Figure S13) (29). We found that when comparing farms within the same landscape, smaller farms accounted for a greater percentage (~40% for farms < 2 ha) of crop species richness compared to larger farms (Figure 5).

Despite the limited studies on crop diversity and farm size, the literature suggests important context-dependencies to the farm size and crop diversity relationship. For example, smallholder farmers engaged in subsistence systems planted a greater diversity of crops to meet their own nutritional needs (49). In market-based systems, smaller farms engaged in niche/specialty markets would sell traditional varieties and species; but often, these same farmers would also grow "modern" varieties for bulk markets that further diversified their production (50). Smaller farms also employed crop diversification to mitigate drought risk when favorable growing conditions outweighed financial limitations (107).

However, there were contexts where larger farms were more diverse. In higher income countries, larger farms were able to more easily enter new markets and they were more willing to diversify production to mitigate potential risk even when it meant lower short-term profits per farm (48, 68). In regions where the average farm size is smaller, there may be minimum size thresholds that prevent crop diversification (48): Smaller, subsistence-based farms were not always able to produce a diversity of crops due to varietal constraints, input availability, and land constraints (66, 67). In Ethiopia, for example, Teshome et al. found sorghum varietal diversity was lower on smaller, subsistence-based farms because farmers were not able to save/obtain enough seed per variety for the subsequent year's planting. Their study also found that the available cropping area was often limiting for multiple varieties on smaller farms since sorghum is wind-pollinated and requires some minimum spacing between varieties to prevent genetic erosion (66).

3.3.5 Resource-use efficiency and GHG emissions

We found no conclusive relationship between farm size and per crop resource-use efficiency (defined by technical efficiency (TE)), and a negative but uncertain effect of farm size on GHG emissions. From 34 studies, we estimated that 50% of studies (95% CI = 39 to 58%) find larger farms to be more resource efficient than smaller farms. For the 18 studies we were able to extract regression slopes from, we found that average farm size had no conclusive effect on resource-use efficiency (0% mean effect; 95% CI: -1 to 2%). We did find evidence that GHG emissions per unit output showed a tendency for being lower on smaller farms (-4% mean effect; 95% CI: -10 to 2%), suggesting that small farms might be more efficient per unit output (Figure 10).

Studies of both resource-use efficiency and GHG emissions represent different aspects of farm level resource-use, and use different assessment methods and literatures. GHG emission estimates used in the Life Cycle Analysis (LCA) literature examine farm level input/output ratios in terms of their GHG equivalents. Resource-use efficiency also examines the farm level input/output relationship but in value or volume terms, and enables analysis of the mechanisms influencing efficiency (e.g., farm size). In our dataset, the most commonly included inputs for studies using TE, in order of frequency, were: labor, seed, herbicides, pesticides, fungicides, and inorganic fertilizer. LCA studies included a more diverse array of inputs and outputs than TE studies and often included pre-farm (e.g., fertilizer production, infrastructure construction, machinery use, etc.) and on-farm energy use (108). In sensitivity tests on TE and LCA inputs and outputs included, there were no significant differences in our findings.

A dominant theme in the literature is that a farm's political, economic, and geographic context -- that is, a farm's access to training, credit, machinery, insurance, inputs, markets, and/or subsidies -- determine its resource-use efficiency. In our meta-regressions, the null relationship between resource-use efficiency and farm size remained even when studies controlled for differences in farmer group membership, credit, or extension access (Figure 10), suggesting that these factors alone do not influence the resource-use efficiency to size relationship.

While we could not test for other factors affecting resource-use efficiency variation by farm size, the literature revealed three barriers that smallholders face. First, technologies were more accessible/available to larger farms. Investments in mechanization, such as irrigation and harvesting infrastructure, involve large fixed costs that are often beyond the reach of smallholders and can be used more efficiently on larger farms with larger fields. For example, combines can cover wider swaths with fewer turns on larger fields, to enable higher yields with less labor and more efficient input use -- but are only cost-effective on larger fields (62). In contrast, smallholders encounter barriers to accessing these technologies, may farm part-time, or may be inhibited by fragmented plots that require increased travel time (109). Second, small farms frequently rely on family labor (especially where there are limited off-farm employment opportunities), which results in a trade-off between increased crop production and labor inefficiency. The latter affects farms' resource-use efficiency: smaller farms' higher yields achieved through more family labor did not outweigh their overall labor inefficiency (54, 69, 70). However, farm equipment rentals and/or low external input systems may offset the high labor inefficiencies. Third, many authors observed that increased access to information from extension, advisory services, and higher levels of education were associated with highly efficient farms (69–72). While there were limited studies examining the interaction(s) between access to information and farm size, Külekçi (2010) found that, even after controlling for farm size, a farmer's access to information was key to increasing TE in Turkey (69). The same was true for market access, where larger farms had better access to improved rural infrastructure that fostered increased market involvement (110).

3.3.6 Profit

We found no general relationship between profit per hectare and farm size. The probability of a study finding a positive, negative or null impact of farm size on profit per hectare had wide and overlapping confidence intervals (Figure 9); this null result was consistent when we examined the effect sizes, where profit per hectare showed no relationship with farm size (-2% mean effect; 95% CI: -10 to 5%).

The profit results were the most spatially heterogeneous across all variables examined. For certain smallholder-dominant countries (e.g., India and Ethiopia) we found that smaller farms were more profitable, whereas larger farms were more profitable in countries dominated by large farms (e.g., the United States). This may suggest that smallholders have better access to markets and inputs in a smallholder-dominant system, but conclusions are limited since we only had 15 observations. The literature says little about how the farm size-profit relationship varies by region. Additionally, many studies simply considered economic inputs and outputs, rather than considering food calories or nutritional value, which would have been more inclusive of subsistence-based producers; hence, our findings primarily reflect the profit outcomes of market-oriented farmers.

The literature suggested some common mechanisms to explain the variations in profit by farm size. Larger farms benefited from technology built for those farm scales--irrigation systems, harvesters, and other machinery are often not available for smaller farms, or there are minimum economic size requirements for farms to recover their fixed costs in the technology or other on-farm investments (e.g., canals, land-leveling, etc.) (74, 111). Additionally, larger farms were able to purchase inputs in bulk to save on upfront costs and had more consistent and streamlined market access, where they could leverage the scale of output for better prices (74). Smaller farms, especially in remote and resource-poor communities with poor access to markets, often sold their crops during harvest to intermediaries at the farm-gate at lower prices (75). Smallholders' lower incomes paired with limited market access and off-farm opportunities may critically affect their food security (112), even under circumstances where they produced a diversity of crops for home consumption (22). In countries with land-based subsidy programs (e.g., the U.S. and China), larger farms achieved a higher profit, in part due to having more available land to receive government payments. For example, Rada et al. (2015) found that in China, larger rice and wheat farms that had access to land consolidation incentives (i.e., expansion subsidies that offered payouts when selling to larger family, specialty, or cooperative farms) achieved higher profit per hectare; however, smaller farms showed higher profits for maize, a sector that did not have these incentives (76).

There were several contexts where improved support for smallholders enabled them to have greater profit than large farms. Small farms were more profitable than large farms when they engaged in niche, high-value markets -- for example, export-oriented producers with fair trade or environmental certification(s) in low or middle-income countries (50). Smaller farms may be more competitive when involved in farmer cooperatives that support group input purchases and foster bulk sale of graded product. In Nicaragua, Deininger et al. (2003) found that profits were greater for smaller farms once the credit system stopped being oriented towards large-holders through structural policy reforms in the mid-1990s (73). Despite these structural and contextual changes that can promote the profitability of smaller farms, several authors cautioned that, even in conditions where larger farms may farm less intensively and obtain lower monetary returns per hectare, they may still obtain higher overall profits and thus outcompete smallholders (113). For instance, in India, Gaurav and Mishra (2015) found that even though smaller farms had greater profits per hectare, their absolute levels of production were often too low to maintain viable livelihoods (111).

3.4 Research gaps

We identified several key research gaps in the literature. First, this body of literature largely focused on important production outcomes for farmers, like yield and efficiency. To inform land reform initiatives aimed at bolstering food security and economic development, future research should also examine the relationships between farm size and livelihoods/well-being for farmers, laborers, and/or consumers. For example, smallholders' higher yields but lower absolute levels of production raises questions about the sustainability of their livelihoods (111, 114). A better understanding of the relationship between farm size and overall agricultural and non-agricultural gross national product (GNP) (or other national indicators) may inform cross-sector economic development planning efforts, such as identifying rural off-farm sectors to promote supplemental incomes for smallholders. Additionally, the IR was investigated in terms of either kg/ha or output value/ha, but recent evidence suggests that smaller farms grow an important array of crops for nutrient diversity (20, 29). While the relationship between macro and micro-nutrient production by farm size is beginning to be understood (20, 29), it is not known who is consuming these nutrients.

Second, there were a limited number of studies on how production methods varied by farm size and the associated environmental implications. While resource-use efficiency metrics look across all inputs to characterize the system as a whole, there is limited detail on particular inputs and farming practices (e.g., do smaller farms use less input-intensive production methods?) and how institutional mechanisms (e.g., credit or farmer groups) influence different farm sizes' access to inputs and information. The LCA literature provides greater detail on inputs and farming practices, but these studies did not often compare across farm sizes. Assessing if particular farm sizes are associated with a set of production methods within and across different cropping systems would enable better identification of scale-specific relationships between farm size and environmental impacts.

Finally, few studies to our knowledge have considered how other social variables and cultural services provided by agriculture vary by farm size. Historically, the social relationships developed and maintained on smaller-scale family farms -- between family members, neighbors, peers, and also between growers and eaters -- provided opportunities to produce and reproduce (or renegotiate) cultural norms, morals, values, and local ecological knowledge (24, 115). Small farms may be more likely to source directly to consumers at farmers markets or through cooperatives, strengthening place-based socioecological linkages. As producers industrialize and scale up, and as rural out-migration continues (with young people in particular leaving agriculture), producers increasingly need to rely on mechanization and hired -- often exploited -- labor (116). The question of labor demand and quality has received little attention in empirical work on both farm size and management practices (117) due to the precariousness of farm work and vulnerability of farm workers (118). Future research should further investigate how issues of labor, health, and wellbeing for laborers, farmers, and consumers interact with farm size and with other sustainability outcomes.

3.5 Conclusion

Our synthesis leverages the past 50 years of empirical evidence that directly measured production, environmental, and economic outcomes of farm size across diverse economic and geographic contexts. We found that smaller farms have greater yields and support higher crop and non-crop diversity, whereas there was no strong relationship between farm size and resource-use efficiency or profit per hectare, but there was a small effect that smaller farms had lower GHG emissions per crop. Additionally, we identified several research gaps. In particular, there were limited cross-regional and cross-discipline studies and many studies of social outcomes (e.g., food security and production resilience) did not compare across farm sizes (e.g., studies on smallholders did not include the counterfactual of larger farms).

Our results suggest that there are multiple trade-offs to consider when assessing the impact of farm size; often, the context in which farms exist is key to understanding the socioecological effects of farm-size structure and farm size transitions. In addition, while it is useful to understand the directionality of the relationship between farm size and each variable in a given context, we propose that it is perhaps more critical from a policy standpoint to understand the mechanism(s) behind each relationship. While the findings for yield and biodiversity conservation generally support the need for investing in small farms, understanding the drivers behind output, earnings, and the other assessed relationships can inform policy opportunities for targeting resources to support the specific processes that lead to multiple desired outcomes across farm sizes.

Finally, our results highlight that in order to support sustainable transitions in farm sizes, more evidence-based synthesis is needed at broad regional scales. Until recently, the role of farm size in the global food system has largely been assessed by independent case studies. As international commitments (e.g., SDGs and COP21 INDCs) begin to evolve into actionable funding plans and as countries continue to decide upon land use policies that directly affect the size of farms, it is critical to identify under what contexts farm size affects different socio-ecological outcomes.

3.6 Data and Methods

A systematic review was conducted with reference to the PRISMA guidelines (61). We searched the Web of Science and Scopus databases for studies in English published prior to December 2017. We used four inclusion criteria: 1) peer-reviewed; 2) directly measured farm size and the outcome variable(s) of interest; 3) reported error estimates/significance tests in determining effect size; and 4) compared farms with similar management systems (e.g., compared small and large maize farms, not small vegetable farms to large cereal farms). Our search yielded 1474 studies. In total, we identified 118 studies (318 observations) that met our criteria; from these, we coded 111 studies (218 observations) as vote counts, of which we extracted regression coefficients from 45 studies (70 observations).

Studies were coded at the observational level to analyze multiple crops, years, and locations per study; studies had multiple observations if they separately reported different crops, years, and/or locations per outcome variable. The main conclusions were categorically coded as “vote-counts”, where an increase in farm size was associated with a decrease, increase, or null relationship to the variable of interest (we found no non-linear results in the literature). For yield, resource-use efficiency, and profit we extracted several additional variables to calculate pooled effect sizes of regression model coefficients. Due to finding a limited number of studies that directly measured farm size and GHG emissions per unit output, we leveraged the Clark and Tilman (2017) meta-analysis database containing 742 agricultural life-cycle analysis (LCA) observations from 152 unique studies (108); we coded observations that reported average farm size to construct a dataset containing crop species, GHG emissions per unit output (in CO_2 equivalents), average farm size, and sample size for 100 observations (11 studies) that met our inclusion criteria.

3.6.1 Synthesis of results

We ran three types of meta-regressions to synthesize the vote count findings, extracted regression slopes, and the GHG emission estimates. First, we used cumulative link multilevel models (CLMM) to synthesize the ordinal vote count findings for yield, resource-use efficiency, profit, and biodiversity (119, 120). We used CLMMs to examine the probability of the ordinal outcome variable (observation finding negative, null, or positive relationships with farm size). For all models, we set the study as a random effect. Because there were more observations available for yields and non-crop biodiversity, we also set crop type as random effects. For non-crop biodiversity, we also set non-crop species type as a random effect. We tested if the additional random effects used for yields and non-crop biodiversity changed the results compared to using only studies as random effects and found no differences.

Second, we used random effects meta-regressions of the standardized regression slopes and standard errors (121, 122) to calculate pooled effects for yield, resource-use efficiency, and profit. Since certain variables contained multiple currencies, efficiency units, or measurement metrics, we relied on Rodríguez-Barranco et al.'s technique to convert farm size regression coefficients and standard errors into standardized regression coefficients (34, p 4, Table 1) (123). Our standardized coefficients represent a relative change in the outcome variable per 1 ha change in farm size. We used the same random effects per variable as in the CLMM models. Sensitivity tests were conducted through cumulative meta-regressions for continuous variables (e.g., year of study, average farm size study observed, etc.) and subsetted meta-regression for categorical variables (e.g., type of diversity metric used, if yield was defined by weight/ha or value/ha, if resource efficiency was derived from data envelopment analysis or stochastic frontier, etc.). All sensitivity tests found no differences in results. Forest plots are given in Figure S14-S16. An inclusion of bias analysis was conducted through funnel plots that compare the observed outcomes to standard errors. There were no clear biases for yields and resource efficiency, but a slight positive bias for profit (Figure S17).

This meta-regression framework also enabled us to further test if the variation in findings between different studies could be attributed to the inclusion/omission of variables that authors used when estimating the relationship between farm size and the variable of interest, through sensitivity analyses using moderators. For yield, we assessed the importance of moderators such as the types of production methods, institutional characteristics (i.e., credit markets and access, extension access, and involvement in farmer cooperatives), and types of labor (i.e., general labor market imperfections, family labor, and household size). Our logic was that if the relationship is moderated by these factors (e.g., if the main relationship became null) it would indicate that there is a systematic variable omission bias in the literature that, once corrected for, could explain the inverse farm size to yield relationship. For resource-use efficiency, we conducted a similar sensitivity analyses, by including moderators that described development interventions (i.e., credit access, extension access, or farmer group membership). Our key hypothesis was that having similar access to credit, extension, or inputs and markets (through farmer groups) may enable small farms to be equally or more efficient than large farms.

Third, for the GHG emission observations, we used robust linear mixed-effects models where we set location and crop type as random effects. To predict GHG emissions per unit output, we used the log average farm size of a study as a fixed effect. The key difference in the GHG emission model is that the data is at the aggregated farm level, as opposed to extracted regression coefficients for the yield, resource-use efficiency, and profit models. (Formulas and further detail on each meta-regression used is available in Appendix B).

Chapter 4: Examining variations in economic productivity and incomes for small-scale farmers

4.1 Abstract

It has long been established that smaller farms have higher yields than larger farms across a wide range of crops and geographical contexts (see Chapter 3). But there has been less empirical evidence on whether there are scale constraints on economic returns to income that prevent small-scale farmers from rising above national poverty lines. We used the Rural Livelihoods Information System (RuLIS) harmonized micro-dataset to examine farm size, productivity (as profit per ha), and income (as profit per person living in an agricultural household) across 34 countries. Our results confirm that smaller farms are more productive but have lower incomes than larger farms. For 30 countries, we used mixed models and scenarios to examine how these relationships varied across common definitions of small-scale farms (i.e., farm size in hectares, the country relevant farm size, the country relevant economic size, the percent of family labor used, and the farming household's level of subsistence). We further tested how the proposed SDG 2.3 definition of small-scale farms, which combines relative farm size with relevant economic size, maps onto poverty. We find that the proposed definition will identify the greatest proportion of farmers living below national poverty lines in a country compared to the alternative small-scale farm definitions tested. Yet, we caution that for SDG 2.3 this definition will need to be disaggregated to ensure other disenfranchised farmer groups (e.g., women, indigenous peoples, certain castes, etc.) are not

neglected. In our country-level scenario analysis, we find that the SDG 2.3 target of doubling the incomes of small-scale farmers will result in an average of 60% of farmers over their national poverty lines regardless of the definition used. Hence, the SDG 2.3 target will need to be modified to support extremely impoverished farmers.

4.2 Introduction

Small-scale farmers are amongst the poorest populations in many lower-income countries (6, 7). International policies, such as the Sustainable Development Goals (SDG), call for supporting small-scale farmers in order to combat poverty. A key target of SDG 2 (Target 2.3) is to "double the agricultural [labor] productivity and incomes of small-scale food producers, in particular women, indigenous peoples, family farmers, pastoralists and fishers." This target is further contextualized against the backdrop of SDG 1, which aims to end poverty, and SDG 2.4's target of sustainably increasing land productivity (37). To support these goals and refine appropriate targets, decision makers and civil society organizations need a better understanding of the interplay between small-scale farmers, productivity, profit, and poverty.

SDG 2 present productivity and incomes as dually achievable objectives, but doubling farmers' land and/or labor productivity may not sufficiently raise farmers' incomes above national poverty lines. For example, there has been a recent upsurge in research promoting intensification and labor-saving technologies for small-scale farmers (124–126). Yet, intensification and labor-saving technologies has not been found to transition farmers out of poverty (125). In 2019, Harris used data from 15 sub-Saharan African countries to examine increases in income following the adoption of intensification technologies and practices (125). In his sample, he found that an increase in agricultural returns from current levels of ~$80-$400/ha/year (2005 USD PPP) to $2500/ha/year would be required to transition 50% of farmers out of poverty. Given that smaller farms have higher yields than larger farms across a wide range of crops and geographical contexts (127), these levels of intensification gains -- even if paired with increased access to cheaper inputs or more integrated supply chains -- are unrealistic for most small-scale farmers given their current land constraints (128). Hence, doubling small-scale farmers' incomes might not be a high enough target to transition many farmers out of poverty.

Beyond individual case studies on the relationship between farm size, profit, and income (111, 114, 129–131), which are sparse when compared to the widely investigated farm size-to-yield relationship, the questions remain: Are there scale constraints on returns to income that prevent small-scale farmers from rising above national poverty lines? Will doubling small-scale farmers' incomes be enough to transition farmers out of poverty? What types of small-scale farmers can transition out of poverty when doubling their incomes and what farmers will require a more aggressive target?

To answer these questions, the operational definitions of small-scale farms that the SDGs can rely upon to identify impoverished farmers need to be further examined. Farm size has become a predominant definition for small-scale farms because it is easy to define and is often captured in sample surveys and agricultural censuses. Recently, earth observation has streamlined field size classification (21, 89, 90), which currently is being linked to farm size (27). Conventionally, farmers operating under 2 ha are defined as smallholders (a synonym for small-scale farmers). This break point is arbitrary, lacks a country relative perspective (e.g., by this definition, Brazil would have 20% of its farms classified as smallholders, while India would have ~80% of its farms be classified as smallholders), and does not account for other key dimensions of impoverished farmers (e.g., farms' economic size, labor dimensions, and market orientation) (1, 12). In addition, global statistics on the role of small-scale farms in the food system are obfuscated by differing definitions, such as, "small farms" (1, 19, 20, 29), "peasant farms" (30), and "family farms" (11, 18).

There are multiple definitions of small-scale farms that differ from spatial farm size. For example, the recent recognition of family farms has gained support for farmers who have long-term relationships to their land and communities (11, 18, 132). Many countries have policy supports for family farms based on different legal definitions, such as the farm's corporation status, level of family labor, farm size thresholds, or a combination of these factors (18). Yet, a farm's corporation status may not reflect desired management practices or a level of family labor that is beneficial to the land, farmers' incomes, or local communities. The literature leading to and stemming from the United Nations' (UN) 2015 "International Year of the Family Farm" reflected this widely diverse definition of "family farms" that ranged from cooperation status, level of family labor, farm residency, farm size, and source of income (12, 18, 132, 133). Another key definition of smallholders is their market orientation. Whether a farmer engages in subsistence-based agriculture or is predominately producing for markets, results in different development interventions/types of support. Underlying each definition of small-scale farms are gender disparities, which is a key cross-cutting theme in SDG 2 since there are significant gendered income and skill gaps that need to be addressed as agriculture feminizes, where an estimated 70-

80% of agricultural work in LMICs are performed by women and this percentage is expected to increase (134).

While no single definition of smallholders will encompass all impoverished farmers, international initiatives to improve the welfare of smallholders need to establish an operational definition that can be compared across countries and time. To monitor SDG 2.3, a combined, country relevant definition of small-scale farms has been proposed -- as a country's smallest 40% farms (as defined by spatial size) that are also within the bottom 40% of economic size (as defined in agricultural revenue) (28). While this definition addresses multiple key debates around measuring farm size, there has been little empirical testing to understand the following questions: Will these two dimensions of small-scale farms identify the most impoverished farmers? Is this indicator a suitable proxy for other key definitions of small-scale farms and what breakpoints capture low-income farmers? Do family farms, subsistence levels, or other definitions of small-scale farms capture a different farming population than farm size and economic size? Will these operational definitions also identify female headed households and farms with female dominant labor?

4.3 Objectives

This study has three objectives. First, we examine the relationship between farm size, land productivity (in profit per ha terms), and income (in profit per all people living in a farming household terms) across 34 countries, using the Rural Livelihoods Information System (RuLIS). The goal of this analysis is to test if farms with high productivity have higher incomes.

Second, we test whether the productivity and income relationships to farm size are better explained by other definitions of small-scale farms (i.e., family labor, country relative economic size, country relative farm size, and subsistence levels) across 30 countries using RuLIS. By empirically testing different definitions of small-scale farms, this analysis seeks to build upon the critical social research that has challenged simplified definitions of smallholders (12, 18, 132, 133).

Third, we examine which types of small-scale farmers (e.g., different small-scale farm definitions and different points on the farm size or income distribution scales) can transition out of poverty if SDG 2.3 target in doubling their incomes was achieved. The goal of this analysis is threefold. We test if doubling farmers' incomes is an appropriate target to transition farmers out of poverty. Then, we test the efficacy of the proposed SDG 2.3 definition of small-scale farms in targeting impoverished farmers. Finally, we provide different thresholds for defining small-scale farms (e.g., 40% smallest farms versus 60% smallest farms in a country).

4.4 Data

RuLIS is a harmonized dataset across 37 countries that uses nationally representative sample surveys. Data is available at the micro level (i.e., individual household records) and contains over 200 harmonized variables on agricultural households' on and off-farm revenues and expenditures. For this analysis, we used two subsets of 34 and 30 countries based on the availability of variables (Figure S18). While RuLIS contains panel datasets for a selection of countries, we used the most recent year per country because we were interested in the broad relationships between farm size, profit, and different smallholder definitions across regional contexts. Summary statistics are presented in Table 4 and Spearman rank correlations are in Figure 11.

Table 4: Summary statistics for data included in analysis.

Small-scale farm definitions

Statistic	No. Obs.	Mean	St. Dev.	Min	Pctl(25)	Pctl(75)	Max
Farm Size (rel)	38,126	0.29	0.25	0.00	0.08	0.45	1.00
Farm Size (ha)	38,126	2.54	3.70	0.25	0.75	3.00	162.00
Female Labor	38,114	0.51	0.22	0.00	0.40	0.62	1.00
Family Labor	38,126	0.97	0.05	0	1.0	1	1
Subsistence	38,126	0.24	0.28	0.00	0.02	0.40	1.00
Economic Size (rel)	38,126	0.28	0.27	0.00	0.05	0.46	1.00

Profit per ha (USD 2011 PPP)

Statistic	No. Obs.	Mean	St. Dev.	Min	Pctl(25)	Pctl(75)	Max
Crop	37,683	1,490	2,547	-11,885	254	1,738	53,524
Livestock	37,062	720	1,654	-10,603	0	646	27,111
On-Farm	36,647	2,202	3,266	-11,885	408	2,632	53,524
Total	38,126	4,384	6,031	-20,409	888	5,408	90,590

Profit per capita (USD 2011 PPP)

Statistic	No. Obs.	Mean	St. Dev.	Min	Pctl(25)	Pctl(75)	Max
Crop	37,717	1.14	1.71	-8.84	0.20	1.43	23.40
Livestock	37,512	0.45	0.89	-3.40	0.00	0.56	12.60
On-Farm	36,954	1.59	2.02	-7.84	0.36	2.10	25.30
Total	38,282	3.07	3.53	-18.60	0.80	4.02	43.40

Since the key policy interest in small-scale farmers is to support the most vulnerable farmers, we tested how the different definitions of smallholders explain on-farm productivity, on-farm income, and total income. Specifically, we looked at the following definitions of small-scale farms: actual farm size (ha), the relevant farm size (country level ranked percentiles), the relevant economic size (country level ranked percentiles), the farming household's level of subsistence, and the percent of family labor a given farm used compared to all labor on the farm (definitions available in Table S11). RuLIS contains data on the number of days family members worked and the total labor days available for only nine countries. To expand our results to include more countries, we used a proxy for the percent family labor -- one hundred minus the percent of labor expenditure compared to the total on-farm expenditures -- that assumed family labor was unpaid; we validated this assumption with sensitivity tests shown in Figure S19. We examined how the SDG 2's cross-cutting theme of gender related to the small-scale farm definitions and tested if it moderated the other smallholder definitions' relationships with productivity and income, where female labor was defined as the percent of female labor used on a given farm compared to all labor on the farm.

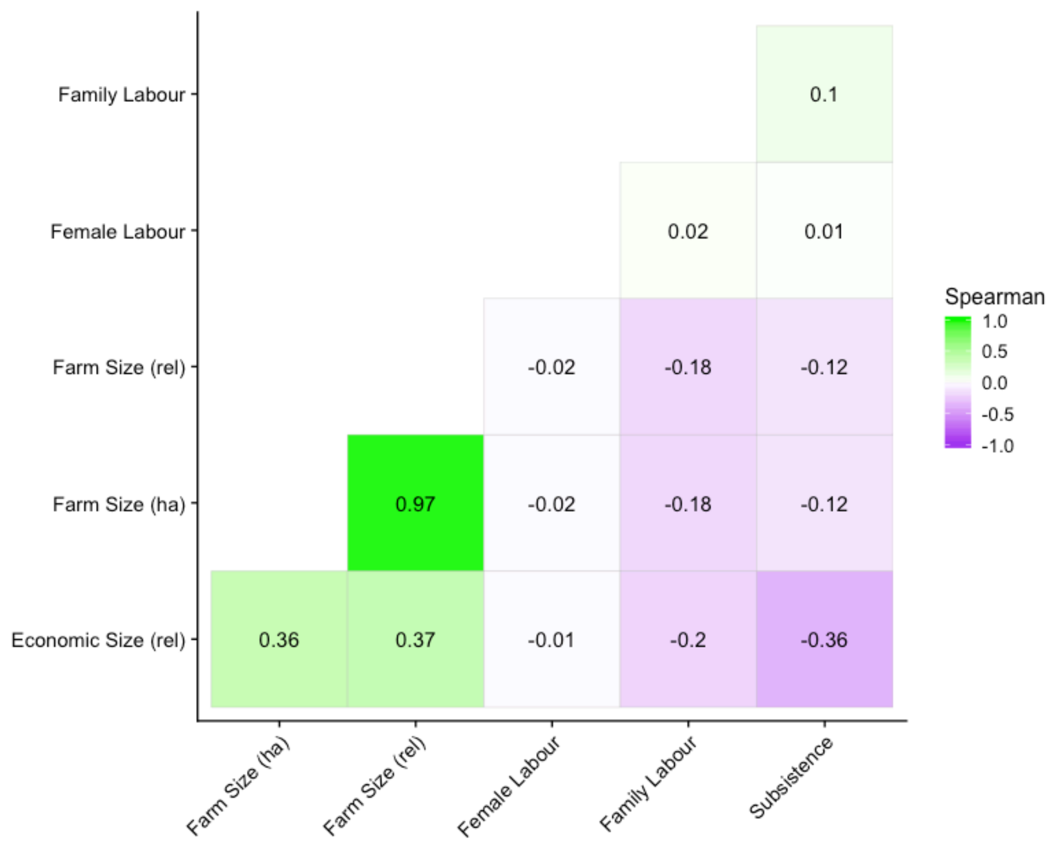

Figure 11: Pooled within country Spearman rank correlations comparing definitions of smallholders. Green represents positive correlations and purple represents negative correlations. All data was used from either subset A or subset B depending on variables' availability.

Since RuLIS consists of nationally representative sample surveys, we restricted the data to only include agricultural households. RuLIS categorizes agricultural households based on whether their share of income from agricultural activities (i.e., crop, livestock, forestry, fishery, and agricultural wage) is higher than 30% or for which a negative share is not due to negative income in agriculture but to a high negative value in income due to non-agricultural and self-employment activities.

To measure farm size, we used a combination of farm size variables since there were different farm size variables available per country. In order of availability we used: cropland, cultivated area, and operated farm area. These farm size variables were used to be consistent with the operational definition of small-scale farms to monitor SDG 2.3 (28).

We defined productivity as profit per ha and income as profit per all people living in the farming household per day, where profit was defined as revenue minus operation expenses. We conducted analyses for on-farm income and for total income (i.e., on and off-farm income) separately; due to similar findings, we only present total income as it is more indicative of a household's entire income. For all analyses, food produced for household consumption was not included as revenue. Instead, we directly tested if food produced for household consumption had an effect on productivity and income by including farms' levels of subsistence as one of the small-scale farm definitions. Similarly, we did not quantify and incorporate unpaid family labor into expenses, as we directly tested if unpaid family labor had an effect on productivity and incomes. In the next section, we detail that all small-scale farm definitions were used as dependent variables to examine their relationships to productivity and incomes.

We analyzed revenues and expenses in real terms (i.e., 2011 USD purchasing power parities (PPP)) and country relative terms (i.e., country level ranked percentiles). We explicitly examined differences between real terms and country relevant terms to address critiques on the true comparability of PPP values (135) and to inform the last stage of our analysis where we compared country defined poverty lines to the smallholder definitions; since models using the real terms were prone to greater heteroskedasticity than models using the country relevant terms, we presented the models with the country-relevant terms. Because only three countries included forestry and fishery variables, we defined on-farm productivity and incomes using livestock and crop values; total income was already calculated in RuLIS to include all on and off-farm sources. While the omission of forestry and fisheries is a limitation to our conclusions, they accounted for ~16% and ~5% of farms' on-farm income in countries reporting these data had income from these sources and we found no relationship between fishery or forestry income and farm size.

4.5 Methods

Our methods were twofold. First, we built weighted hierarchical mixed effect models to investigate the relationships between the small-scale farm definitions, productivity, and income, and whether these relationships held when controlling for other small-scale farm definitions.

$$y_i = W_i X_i \beta + Z_i b_i + \varepsilon_i$$

In the above equation, y is the standardized outcome variable for each household, *i* (136). $X_i\beta$ is a matrix of fixed effects, which include each definition of small-scale farm. W_i are a vector of standardized survey sampling weights, where each national survey dataset includes weights to adjust national representativeness based on the households sampled and the known population; since each survey presents their sampling weights according to their known populations, we standardized each country's sampling weights by dividing each weight by the maximum weight per country to ensure values are between 0 to 1. All other continuous variables were centered around their mean and standardized by their standard deviation ($\frac{x-\bar{x}}{\sigma}$) in order to directly compare the direction and magnitude of the independent variables. Z_ib_i are the random intercepts, which are the primary sampling unit (PSU) of a survey nested within the country the survey was administered to account for survey sampling differences, survey representativeness, national and local policies, comparison between similar local geographies, and other possible local contextual factors (137); the $X_i\beta$ matrix includes a pooled intercept derived from the Z_ib_i matrix. ε_i is the error term. We bootstrapped the models to extract the fixed effect intercept, then calculated the median effect and 95% confidence intervals.

The weighted hierarchical mixed effect models were used to ensure within country correlation structures were accounted, to allow for generalizability of results to countries not in the analysis, and to enable greater efficiency of estimates due to the sharing of information within the model's structure (136, 137). For comparison, we also examined weighted fixed effects models where countries and PSUs were used as fixed terms. We tested if non-weighted regression and robust regressions (i.e., robust mixed models and sandwich estimators) changed our coefficients and error estimates; all models showed similar coefficients and standard errors. In the models that contained all small-scale farmer definitions, we did not include farm size in relative and hectare terms due to high multi-collinearity; instead, we constructed separate models. For the fully parameterized, hierarchical mixed effects models, we tested the amount of variance that each small-scale farm definition explained through bootstrapping the analysis of variance (ANOVA) test; boxplots were used to compare median and distributions of these results.

In the RuLIS dataset, countries have different available variables. To include the most countries and to test the most definitions of small-scale farms, we conducted two different sets of regressions. Set A used 34 countries and consisted of farm size (in relative and hectare terms), relative economic size, and percent female labor. Set B used 30 countries and consisted of all the variables in Set A, but also percent family labor and percent subsistence (see Figure S18 for spatial distribution). The outcome variables for both sets remained the same. Since our results were consistent across both subgroups and Set B allowed us to test more definitions of small-scale farms, we present Set B as our main results.

For the second stage of our analysis, we conducted a simple set of scenarios to understand the percentage of farmers who transition out of poverty by doubling their total incomes. These scenarios began with doubling all farms' total incomes and classifying farms into two groups per country: those that were above their national poverty line (Group A) and those that were below their national poverty line even after doubling the income (Group B). We then used the surveys' sample weights and expansion factors to calculate the national number of farmers in each group. We then calculated the percentage of farmers who would be out of poverty after doubling their incomes. Then, we compared these results to the percentage of farmers that have incomes over the national poverty lines with a different tuning parameter (e.g., two versus three, or four times their incomes; one times their incomes represents the present-day percentage of farmers who live above their national poverty lines). For all farms that had equal to or less than $0 USD PPP incomes, we grouped them into Group B as doubling zeros and negative incomes as there were less than 3% of these observations in our dataset. The analysis was conducted at the country level and 95% confidence intervals were computed to understand the variation across countries.

By expanding these scenarios, we further tested how doubling certain groups of farmers' incomes related to farmers transitioning out of poverty. We used each of the key definitions of small-scale farms to calculate the percentage of farmers that are out of poverty when their incomes were doubled and with different thresholds for defining smallholders (i.e., doubling the incomes of farmers on the smallest 40% of farms per country or farmers living in households with a female dominant labor force). Since thresholds for farm size in ha terms are operationally and theoretically difficult to compare across countries due to the farm size distributions across different regions (e.g., in Asia, ~80% of farms are < 2 ha, while in Latin America and the Caribbean ~25% of farms are < 2 ha), we focused on relative farm size. We plotted a range of thresholds to test the appropriate cut-off point to have the most farmers transition out of poverty (e.g., will doubling the incomes of farmers on the smallest 40% of farms get 100% of farmers out of poverty or should we target the smallest 60% of farms?). Since each definition of smallholder captured different types of farms and farmers, we then combined the key definitions of smallholders and plotted different thresholds to understand which combination and threshold would result in the highest number of farmers transitioning out of poverty when their incomes are doubled (e.g., the smallest 40% of farms per country and farms with the lowest 40% of revenues per country).

4.6 Results

Results from the hierarchical mixed models showed that smaller farms had higher productivity (slope: -0.54; 95% CI: -0.56 to -0.53), yet lower incomes (slope: 0.18; 95% CI: 0.16 to 0.19) than larger farms. These results held when all other definitions of small-scale farms were controlled except for economic size, which had a moderating effect on farm size's income relationship (Table

5). The country relevant economic size of a farm showed consistent positive relationship with productivity and income in all models, i.e., smaller farms (on economic basis) were less productive (slope: 0.10; 95% CI: 0.09 to 0.11) and had lower incomes (slope: 0.36; 95% CI: 0.35 to 0.37) than larger farms. There were moderate negative effects between family labor and productivity (slope: 0.13; 95% CI: 0.12 to 0.14) and moderate positive effects between family labor and income (slope: -0.08; 95% CI: -0.06 to -0.09). Farms' level of subsistence showed moderate negative effects with productivity (slope: -0.02; 95% CI: -0.01 to -0.03) and income (slope: -0.21; 95% CI: -0.22 to -0.19). There were no trends with female labor, productivity, or incomes. Fixed effects models showed relatively consistent results to the hierarchical mixed models (Table S12-S13). **Error! Reference source not found.** shows the predicted relationships, 95% confidence intervals, and raw data for the fully parameterized models (i.e., models RE 10 and RE 20). Country subsets that used 30 or 34 countries showed similar results for the overlapping variables. Other sensitivity tests were conducted on these results to compare robust methods on random samples of the data (because the complete dataset was too large to compute computationally intensive robust mixed models) -- no large inferential differences were detected. Here, we highlight the relationships between country relevant total profit and the dependent variables, but we also found consistent relationships when using country relevant on-farm profit.

Table 5: Hierarchical mixed effects models predicting on-farm profit per person living in the farming household in country relative terms. Coefficients and bootstrapped 95% confidence intervals (in parentheses). Models contain the 30 countries in subgroup 2 (Figure S18).

DV	RE1	RE2	RE3	RE4	RE5	RE6	RE7	RE8	RE9	RE10
Intercept	0.04 (-0.06; 0.14)	0.03 (-0.04; 0.11)	0.05 (-0.03; 0.13)	0.11* (0.02; 0.20)	-0.03 (-0.10; 0.04)	0.04 (-0.04; 0.12)	0 (-0.07; 0.06)	0 (-0.07; 0.07)	0.10* (0.00; 0.20)	0.09* (0.01; 0.18)
Farm Size (ha)	0.18* (0.16; 0.19)						-0.01 (-0.02; 0.00)		0.14* (0.13; 0.15)	
Farm Size (rel)		0.14* (0.13; 0.15)						-0.01* (-0.02; -0.00)		0.11* (0.10; 0.12)
Family Labor			-0.08* (-0.09; -0.06)				-0.01 (-0.02; 0.00)	-0.01 (-0.02; 0.00)	-0.04* (-0.05; -0.03)	-0.04* (-0.05; -0.03)
Subsistence				-0.21* (-0.22; -0.19)			-0.07* (-0.08; -0.06)	-0.07* (-0.08; -0.06)	-0.18* (-0.19; -0.17)	-0.18* (-0.19; -0.17)
Economic Size					0.36* (0.35; 0.37)		0.34* (0.33; 0.35)	0.34* (0.33; 0.35)		
Female Labor						-0.07* (-0.08; -0.06)	-0.06* (-0.07; -0.05)	-0.06* (-0.07; -0.05)	-0.06* (-0.07; -0.05)	-0.06* (-0.07; -0.05)
Num. obs.	38270	38270	38270	38270	38270	38270	38270	38270	38270	38270
Num. groups: country:psu	7235	7235	7235	7235	7235	7235	7235	7235	7235	7235
Num. groups: country	30	30	30	30	30	30	30	30	30	30

Table 6: Hierarchical mixed effects models predicting on-farm profit per ha in country relative terms. Coefficients and bootstrapped 95% confidence intervals (in parentheses). Models contain the 30 countries in subgroup 2 (Figure S18).

DV	**RE11**	**RE12**	**RE13**	**RE14**	**RE15**	**RE16**	**RE17**	**RE18**	**RE19**	**RE20**
Intercept	0.02 (-0.08; 0.12)	0.04 (-0.04; 0.11)	0 (-0.06; 0.06)	0.02 (-0.04; 0.07)	-0.01 (-0.06; 0.05)	0.01 (-0.04; 0.07)	-0.05 (-0.19; 0.09)	-0.03 (-0.10; 0.04)	0.05 (-0.05; 0.15)	0.06 (-0.01; 0.13)
Farm Size (ha)	-0.54* (-0.56; -0.53)						-0.71* (-0.72; -0.70)		-0.55* (-0.57; -0.54)	
Farm Size (rel)		-0.39* (-0.40; -0.38)						-0.52* (-0.53; -0.51)		-0.40* (-0.41; -0.39)
Family Labor			0.13* (0.12; 0.14)				0.10* (0.09; 0.11)	0.12* (0.10; 0.13)	0.07* (0.06; 0.08)	0.08* (0.07; 0.10)
Subsistence				-0.02* (-0.03; -0.01)			0 (-0.01; 0.01)	0 (-0.01; 0.01)	-0.11* (-0.13; -0.10)	-0.11* (-0.12; -0.09)
Economic Size					0.10* (0.09; 0.11)		0.35* (0.34; 0.36)	0.33* (0.32; 0.34)		
Female Labor						0.01* (0.00; 0.02)	0 (-0.01; 0.01)	0 (-0.01; 0.01)	0 (-0.01; 0.01)	0 (-0.01; 0.01)
Num. obs.	38114	38114	38114	38114	38114	38114	38114	38114	38114	38114
Num. groups: country:psu	7205	7205	7205	7205	7205	7205	7205	7205	7205	7205
Num. groups: country	30	30	30	30	30	30	30	30	30	30

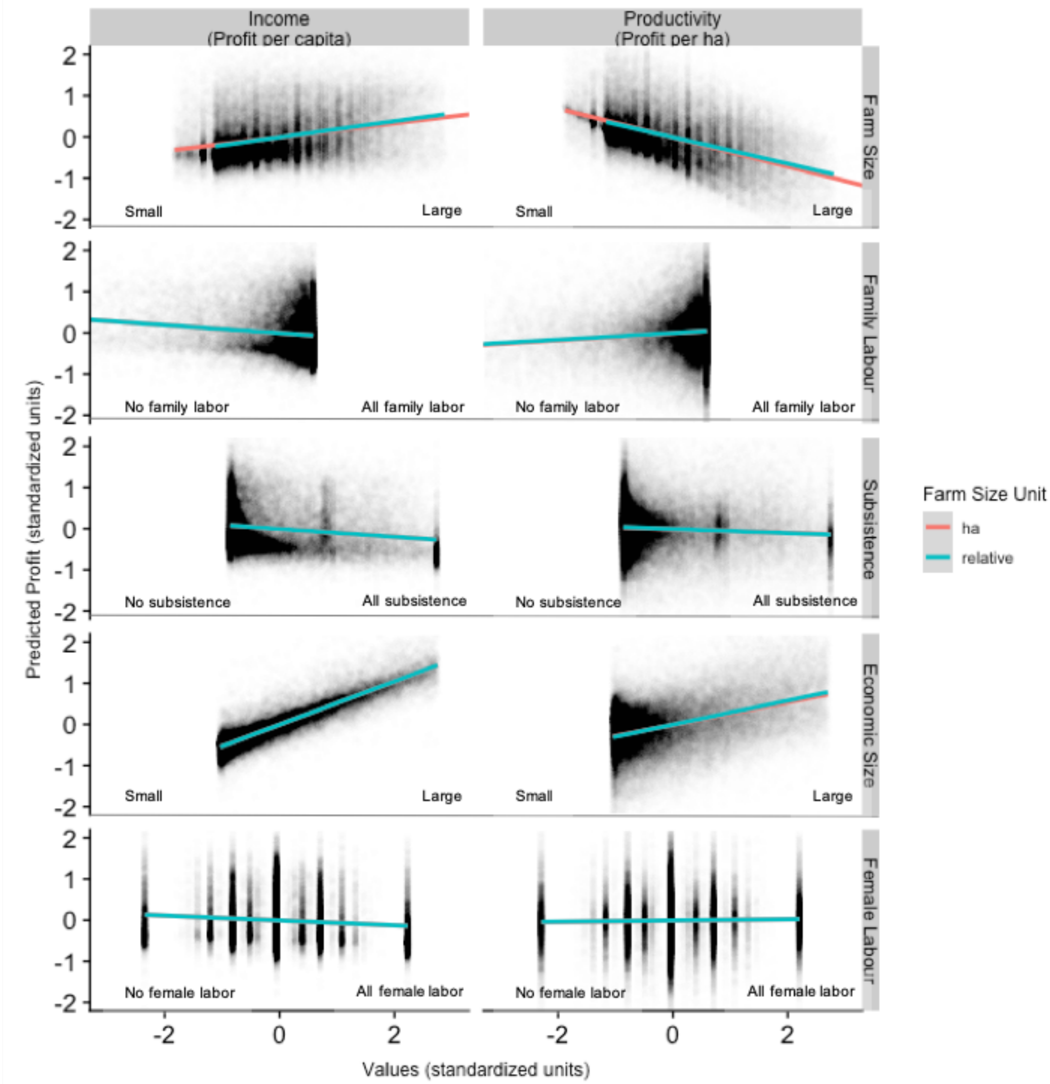

Figure 12: Predicted relationships between smallholder definitions and income (profit per capita - left) and productivity (profit per ha - right). Predictions were based on weighted hierarchical mixed effects models that controlled for all other definitions of smallholders. All variables are in standardized terms. 95% confidence intervals are plotted but are very narrow due to the high sample size. Full data is plotted in black to show the distribution and model fits. The top right and left plots show farm size in ha (orange) and country relative (green) terms. Models contain the 30 countries in subset 2.

In examining the pooled within country Spearman rank correlations, female labor had the lowest absolute correlations with all smallholder definitions tested. The Spearman Rank correlation matrix showed that each definition of smallholder captured a different segment of farm types (Figure 11). Female labor was the least correlated with the other definitions of smallholders (absolute values of the correlations ranged from 0.01 to 0.02), while economic size most correlated with other definitions of smallholders (absolute values of the correlations ranged from to 0.2 to 0.37, not including correlations with female labor).

The bootstrapped ANOVA tests show that farm size (in both relative and hectare terms) explained the largest amount of variance in productivity in both relative (median: 62%; range: 59 to 65%) and hectare (median: 65%; range: 63 to 68%) terms, while country relevant economic size explained the largest amount of variance in income (median: 80%; range: 76 to 84%) (Figure 13). The lower percentages of variance explained by subsistence levels, family labor, and female labor may be due to the heterogeneity of their relationships with productivity and incomes across countries (Figure 14).

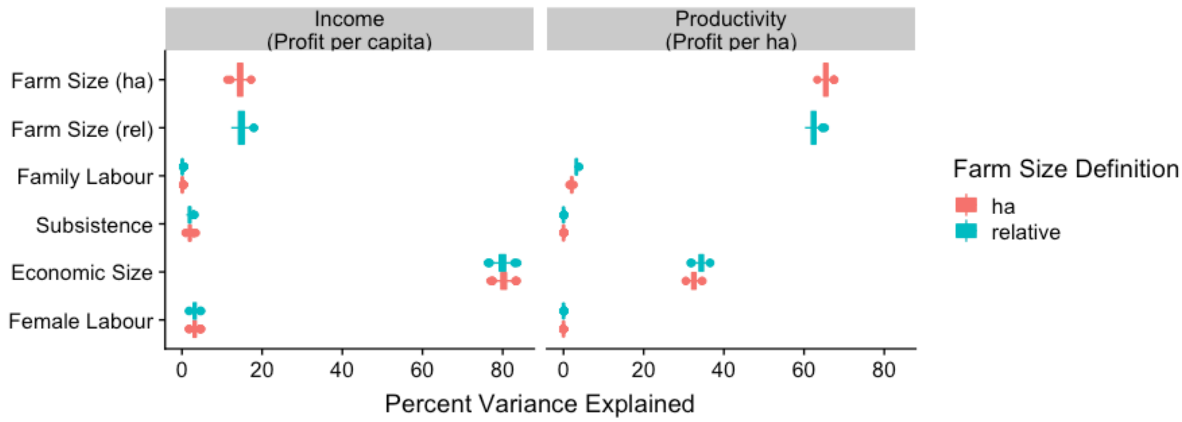

Figure 13: Boxplots show bootstrapped ANOVA results from weighted hierarchical mixed effects models that used all definitions of smallholders. The percent of variance explained per each variable are on the x-axis. Orange indicates that farm size in ha terms was used in the model, while green indicates that farm size in country relevant terms was used in the model. Models contain the 30 countries in subset 2.

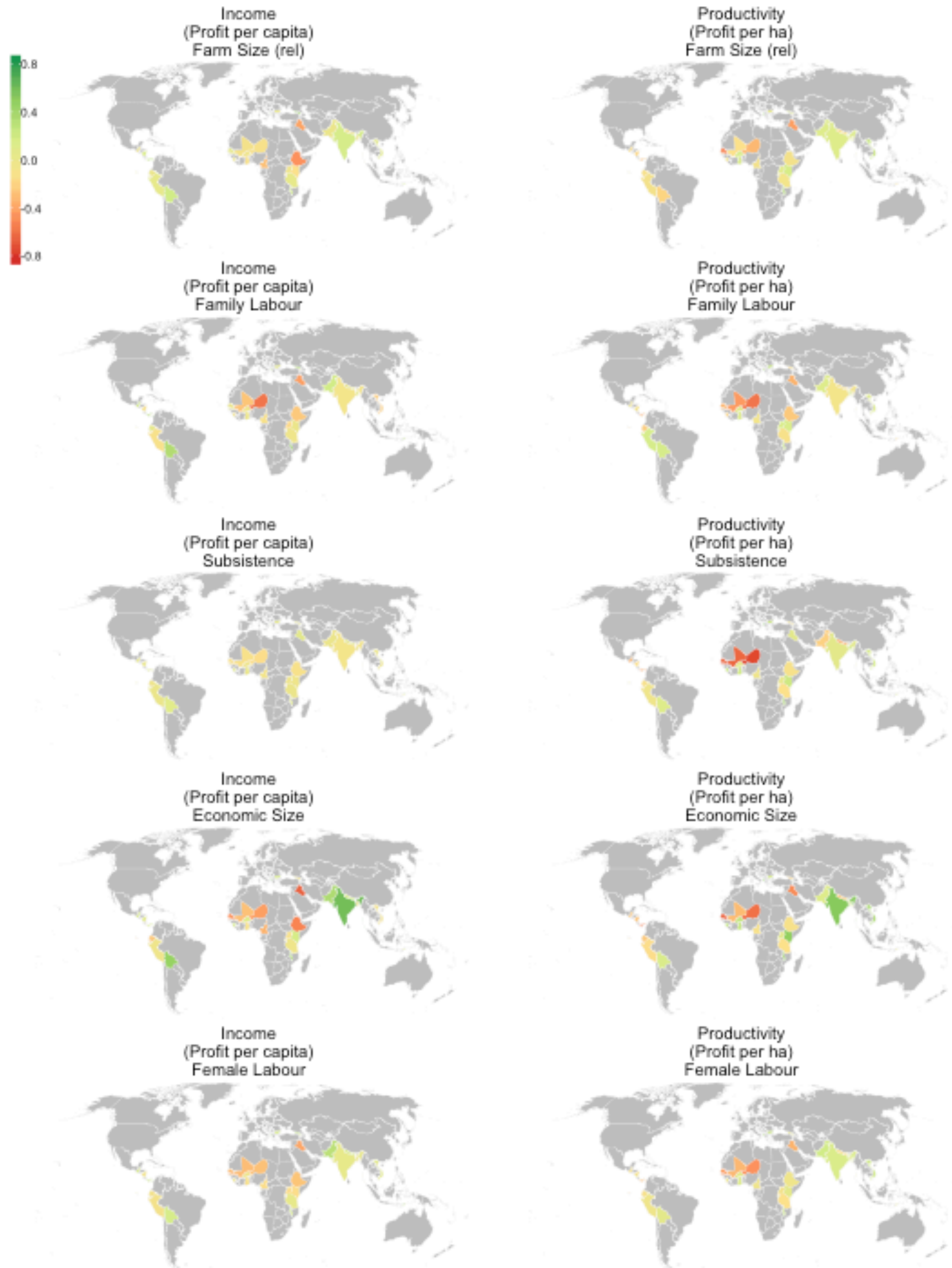

Figure 14: Predicted relationships between smallholder definitions and profit per capita (left) and profit per ha (right). Predictions were based on weighted hierarchical mixed effects models that controlled for all other definitions of smallholders, where the smallholder definition plotted was calculated as a random slope. All variables are in standardized terms. Green indicates a positive relationship between the smallholder definition and profit, while red indicates a negative relationship. Models contain the 30 countries in subset 2.

From our scenarios, we found that across countries an average of 37% (95% CI: 34 to 40%) of farmers were already above their national poverty line (Figure 15). We found that doubling all farmers' incomes would result in an average of 58% (95% CI: 56 to 60%) of farmers with incomes over their national poverty lines. Tripling farmers' incomes would result in 71% (95% CI: 69 to 73%) of farmers over their national poverty lines. Since there was a smaller marginal gain after tripling farmers' incomes (Figure S21) and the SDG 2.3 goal aims to double the incomes of smallholders, we doubled the incomes for remainder of the scenarios.

We found that relative economic size, subsistence, and relative farm size required the lowest threshold (e.g., the lowest 40% of revenue in a country, farmers who consume 40% of their crops, or the smallest 40% of farms in a country) to achieve transitioning the most farmers out of poverty compared to the other definitions of smallholders. Interestingly, once the percentage of female labor was over 50% (i.e., females became a farms' dominant labor source), the variable captured a large number of farmers that could transition out of poverty. Since each definition of smallholder captured different farmers in the population (Figure 11), we combined relative economic size with each of the other identified smallholder definitions that best explained income. Relative farm size, subsistence level, and relative economic size showed similar trends in Figure 15B. We combined relative farm size and subsistence level with relative economic size because relative economic size provided the greatest explanation of incomes' variance by several magnitudes in our regressions (Figure 15C). The relative economic size and relative farm size combination resulted in the highest number of farmers that could transition out of poverty with the lowest thresholds. The combination of relative economic size and subsistence also showed similar trends that required slightly higher thresholds to achieve comparable percentages of farmers transitioning out of poverty.

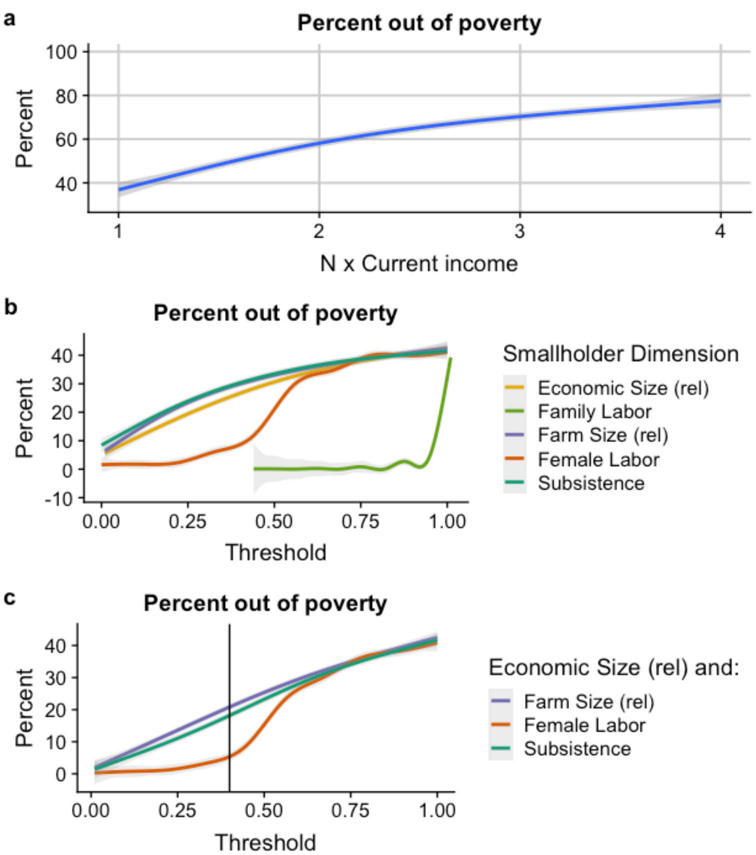

Figure 15: Plot A shows the percent of farmers that have incomes higher than their national poverty lines if all farmers' incomes are increased by a multiple of N (e.g., 1x, 2x, 3x the income). Plot B shows the percent of farmers over their national poverty lines if only smallholders' incomes are doubled according to different thresholds (e.g., if we double the incomes of farmers on farms smaller than 40% of other farms in their country compared to another threshold, such as 60%; another example are for the family farmers, where if we double the incomes of farmers that uses family labor for 50% of all their labor compared to another threshold, such as 80%). The extents of the lines are consistent with available data (e.g., family labor was heavily right skewed so there are few low threshold points). Plot C shows if we combine relative economic size (the best predictor of income from the above regression analysis) with each of the other key definitions of smallholders we identified in the regression analysis: female labor, subsistence, and relative farm size. The 95% confidence intervals in each plot represents cross-country variation.

4.7 Discussion

Our results suggest that, while smaller farms were more productive, households with smaller farms were poorer than households with larger farms. Our findings test the farm size, productivity, and income relationships across the Global South, in which for 34 countries we find consistent results and no clear regional patterns (Figure 14). Our results place the often-observed inverse farm size to productivity relationship into a poverty context, where smaller farms' higher productivity does not, on average, translate into higher incomes compared to larger farms. We found that there was a strong negative relationship between farm size (in both hectare and country relative terms) and productivity (as defined by profit per ha). Yet, there was a near equally strong and positive relationship with on-farm and total income (as defined by profit per person living in a farming household).

Past case studies on the productivity to farm size relationship suggested that unavailability of off-farm opportunities for family members of smaller farms result in higher yields (54, 69, 70, 131) better market access for larger farms (74, 75), and/or recovering fixed operating costs requiring minimal farm sizes (138). By using a subset of the data that included 30 countries, we did not find that family labor moderated our farm size to productivity relationships. While market access and examining discrete costs were beyond the scope of our study, we did find that the relative economic size of a farm better explained the variance in incomes when compared to spatial farm size. Future studies would benefit from determining the greatest costs across farm sizes, if these cost relationships change due to the physical or relevant size of a farm, and if these costs are consistent hurdles across different geographies.

Our results can also be used to interrogate conventional spatial definitions of small-scale farms. We examined how different definitions of small-scale farms can be used to identify impoverished farmer populations across the Global South. Small farm sizes (in hectare terms) has conventionally been used as a proxy for impoverished farmers (19, 20, 29, 81, 83, 112, 131). While we found farm size (in both hectare and country relative terms) to explain a large degree of income, the relative economic size of a farm was a better indicator of farmers' income levels. Family labor, female labor, and levels of subsistence showed moderate relationships with income, but explained a small portion of the variance. Our study tested the critiques on using farm size as a proxy for impoverished farms (12, 18, 132, 133), but we found farm size was still a strong predictor for poverty compared to several other proposed smallholder definitions.

Under our scenarios, doubling farmers' incomes would enable ~60% of farmers per country to have incomes over their national poverty lines and tripling their incomes would enable ~70% of farmers to be over their national poverty lines. These results show that SDG 2.3's aim to double incomes will be insufficient to also meet SDG 1's goal of eradicating extreme poverty for vulnerable groups, such as small-scale farmers. We found that a dual definition of small-scale farms using relative economic size and relative farm size would target the greatest number of farmers to transition out of poverty (as defined by national poverty lines).

Despite our findings, we caution against using reductionist operational definitions of small-scale farms, since no single definition can capture all vulnerable farmers. For example, we found that all definitions of small-scale farms tested were not correlated with female farms, which indicates that the farmers on female dominant farms are most likely to not be identified by the combined economic and farm size definition of small-scale farms. While our analysis does not incorporate trends, our cross-sectional results imply that the feminization of agriculture across the Global South may not be associated with farm size; an idea reflected in past research on the feminization of agriculture. Lastarria-Cornhiels suggested that there may be increased female workers on larger farms due to the general trend across many countries towards export-oriented production (139). At the same time, she illustrated that there is increased formal and informal female labor and ownership on smaller farms as men seek off-farm and/or seasonal employment. While these two types of increased female involvement in agriculture may cancel out any farm size to gender relationships, female farmers still suffer a wage gap from gendered agricultural labor roles (e.g., men assuming jobs using machinery and women skills be deemed as unskilled, such as weeding or grafting) (140, 141). Our finding highlights the importance of treating gender as a cross-cutting theme to achieve SDG 2.3 because the proposed FAO definition of small-scale farm will not identify gender differences. Gender dimensions of agriculture -- as well as religion, caste, and other disenfranchised minority groups that we were not able to test – will need to be disaggregated when monitoring the success of programs intended to inclusively assist small-scale farmers or else they will not be identified with the FAO's proposed definition.

4.8 Conclusion

Our study used data from 30 to 34 countries to generate three findings that can help decision makers and civil society organizations operationalize their commitments to SDG 2.3, which aims to double the productivity and incomes of small-scale farmers. First, we found that smaller farms are more productive (in profit per ha terms) yet have lower incomes (in profit per person living in an agricultural household terms) than larger farms. This result suggests that there are scale constraints on returns to income that prevent small-scale farmers from rising above national poverty lines. Second, we found evidence for the proposed operational definition of small-scale farms, which combines country relevant economic size with country relevant farm size. These two dimensions of small-scale farmers identified more impoverished farmers than the alternative dimensions of small-scale farmers we tested, such as family labor or levels of subsistence. However, none of the definitions that we tested identified poorer female farmers. This finding suggests that gender aspects of small-scale farmers need to be disaggregated to ensure gender inclusion. Finally, we found that the SDG 2.3's doubling target needs to be refined. Our simple scenarios revealed that while some farmers can transition out of poverty when doubling their incomes, a large majority of farmers will require greater support to make the transition.

Chapter 5: How can development policy target smallholders to achieve SDG 2.3?

5.1 Introduction

Sustainable Development Goal (SDG) 2.3 seeks to alleviate rural poverty by doubling the incomes and productivity of small-scale farmers (37). Yet, SDG 2.3 has amongst the least available information in SDG monitoring reports and databases compared to other targets.[2] This data disparity prevents civil society from holding countries accountable to their SDG 2.3 commitments and to plan allocation of public and donor investment in agricultural development policy.

This policy brief outlines two interlinked solutions to address SDG 2.3's data disparity: better defining small-scale farms and harmonizing existing data. In 2018, the Food and Agricultural Organization (FAO) proposed an operational definition for small-scale food producers -- as the smallest 40% of farms and farms with the lowest 40% of agricultural revenue in a given country. Prior to this, there was no consensus on defining small-scale farms, which made leveraging existing data difficult. Recently, several efforts have begun to address SDG 2.3 data needs by combining existing household data that contains socio-economic and agricultural production

2 For example, both the World Bank SDG Atlas (http://datatopics.worldbank.org/sdgatlas/SDG-02-zero-hunger.html) and the UN SDG Indicator Database (https://unstats.un.org/sdgs/indicators/database/) do not have data on target 2.3.

variables. Once harmonized, these data can be linked to geospatial information to enable better monitoring and allocation of resources.

In this policy brief, we present empirical support for the FAO's proposed definition of small-scale farm. Then, we present key considerations needed to make disparate agricultural household datasets interoperable. Through a case study, we illustrate the utility of combining socio-economic and geospatial data sets to identify small-scale farmers' irrigation requirements. To bolster monitoring and planning pathways to achieve SDG 2.3, we conclude with recommendations for improving current data efforts.

5.2 Defining small-scale farms

The FAO recommends that small-scale farms be defined as those that are both the smallest 40% of farms and have the lowest 40% of crop revenue in a given country (28). While a definition that incorporates additional dimensions may be more inclusive of disparate populations, given the current lack of data on additional dimensions, the FAO's definition is the most useful to compare trends across and within countries over time.

By using country relevant terms that combine two key aspects of small-scale farms, the proposed definition addresses many of the critiques of commonly used definitions of small-scale farms. For example, farms operating under 2 ha have been conventionally defined as small-scale farms. Yet, this break point is arbitrary, lacking a country relevant perspective (e.g., by this definition, Brazil would have ~20% of its farms classified as small-scale, while India would have ~80% of its farms be classified as small (1)). By relying on a dual country relevant farm and economic size definition, macro-economic dimensions that affect smaller farms, such as unequal land markets, can be used to identify disadvantaged farmers.

The FAO's proposed definition is not without problems. It does not account for other key dimensions of disadvantaged farmers, such as labor and ownership (e.g., family labor or corporation status), market orientation (e.g., level of subsistence), or geographic considerations (e.g., poor land quality and drought prone areas) (1, 12). In addition, other disadvantaged farmer groups -- such as female farmers, certain castes and religions, and indigenous peoples -- are not included in the definition but these specific groups are listed as key populations in SDG 2.3. These cross-cutting themes, and the political and geographic contexts in which farmers operate, need to be incorporated into the identification and monitoring of disadvantaged farmers. The FAO's proposed definition needs to be disaggregated to avoid misidentifying key groups.

Despite these critiques, we found that FAO's proposed definition for small-scale farms identifies impoverished farmers. We examined how different dimensions of small-scale farms related to rural poverty using the newly harmonized RuLIS dataset across 30 countries (see Chapter 4). Specifically, we examined the relationship between income and several definitions of small-scale farms, including: the actual farm size, country relative farm size, country relevant economic size, percent of family labor used by a farm, and the level of subsistence upon which a farm relied. While these tested definitions do not constitute a comprehensive list of all dimensions characterising small-scale farms, they were available in a common dataset, across a wide range of countries (n = 30), representative at the sub-national level, and available at the individual farm level (RuLIS is a harmonized micro-data). We found that country-relevant economic size and country-relevant farm size explained more variation in farming households' per capita total income (including all on and off-farm revenue and expenses) than the other dimensions of small-scale farms. Smaller farm sizes and smaller economic sizes both had lower incomes than larger farms. There was also a positive relationship between the economic size of farms and productivity (in profit per ha terms) but a negative relationship between spatial farm size and productivity. While SDG 2.3 defines productivity in terms of labor productivity and not profit per ha, our results are consistent with other studies that have found intensification through technology or labor saving techniques does not always correspond to higher incomes (124, 126, 142). These results suggest that the FAO's definition will better identify impoverished farmers than using family ownership or market orientation.

Another key benefit of FAO's proposed definition of small-scale farms is its flexibility and operability. The goal of this definition is to compare countries' progress in reaching the SDG 2.3 target of doubling small-scale farmers' productivity and incomes by 2030. By using country-relevant terms, the same population can be monitored across countries while accounting for within-country changes in agrarian structure over time. For example, if we want to track the change in smallholders' incomes over time, this definition will not be influenced by larger shifts in the overall farm size of a country that can occur in response to economic development, land reform, or land consolidation policies (9). Similarly, when comparing countries that have different types of labor, land, and factor markets, disadvantaged groups of farmers will be consistently identified.

5.3 Available data for SDG 2.3

To address SDG 2.3's data disparity, several existing national sample surveys and agricultural censuses are currently being combined to provide valuable information about the livelihood constraints small-scale farmers face. In 2019, there has been tremendous effort to develop the Rural Livelihood Information System (RuLIS)[3] to harmonize national sample surveys, specifically across variables relating to small-scale farmers. RuLIS represents an ongoing wave of data harmonization needed to better understand the role of small-scale farms in the global food system and to empirically verify if certain interventions have improved incomes and productivity. Other efforts have focussed on discrete questions relating to smallholders such as quantifying the amount of food globally produced by small-scale farms or family farms (18–20, 29, 143).

While newly harmonized data will include disparate years of data collection, the development community plans to coordinate future household surveys with common and interoperable modules around key SDG targets, including small-scale farmers' productivity and incomes. The 50x2030 project is a donor coordinated initiative to bolster timely data collection (conducting multiple sample rounds in 10-year cycles) of agricultural households' production and incomes with standard agricultural survey modules. The project will include three to five additional low and middle income countries (LMICs) a year until 2030, when there will be 50 countries using the same survey modules that build on the World Bank's Living Standards Measurement Survey

3 http://www.fao.org/in-action/rural-livelihoods-dataset-rulis/en/

(LSMS) and FAO's AGRIS. While this project will be critical to long-term monitoring, we still need to harmonize currently available surveys to develop baselines for SDG 2.3.

To bolster monitoring and planning pathways to achieve SDG 2.3, these harmonized datasets can be further combined with geospatial information. With this type of interoperability, one could consider the geographic and environmental contexts in which farmers operate. The combined data could answer key questions such as: What percentage of small-scale farms operate on low quality land? How many small-scale farmers live in drought-prone areas without access to irrigation? How far do they live from a paved road, electricity, cell phone signal, or refrigerated processing/market facility?

Household surveys can be georeferenced by collecting geographic coordinates of interviewed households and/or their fields or by collecting the village, district, or region codes. Two methods can be used to georeference household surveys. First, future surveys can be updated to use field equipment that capture the information (e.g., GPS monitors or smartphones /tablets equipped with GPS monitors). The public health sector has used georeferenced surveys for many years and can provide best practices for data privacy and dissemination for the agricultural sector (e.g., the World Health Organization's Demographic and Health Surveys (DHS)). The second method utilizing national sample surveys does not require additional data to be collected; national sample surveys, which are underutilized, already contain district or region names. There has been a missing processing step to link this information to administrative boundaries (i.e., a shapefile). While this processing step needs to consider the survey sampling designs, many surveys are representative at the subnational scale and a few efforts (see Chapter 2) have already tied such surveys to maps of administrative boundaries (29).

In addition to an improved understanding of the geographic and environmental contexts in which different small-scale farmers operate, linking household surveys with geospatial information can aid agricultural early warning systems. Current early warning systems use market price and climate data in order to identify when a production or economic shock occurs (144). By linking household surveys to geospatial information, these early warning systems can transition from identifying which locations are experiencing production or economic shocks to identifying the populations that are experiencing production shocks. Interventions, thus, can target sub-national regions with small-scale farmers.

5.4 Linking social and environmental data: a case study

To demonstrate how different types of household and geospatial data can be combined for the purpose of improving SDG 2.3 monitoring, we created four global maps at 10km spatial resolution (Figure 16). We use the FAO's proposed definition of small-scale farmer, where small-scale farms are defined as the smallest 40% of farms in a country and the smallest 40% of economic revenue in a country. The farm size data consists of the World Census of Agriculture (WCA) national farm size distributions and a global field size map. In 2015 and 2019, two global field size maps were created by crowd source campaigns, during which volunteers classified satellite images into ranges of small, medium, and large agricultural fields (in absolute terms) (21, 89). While the relationship between field size and farm size is not always linear, this data can be used to downscale the WCA's national farm size distributions (1) to produce a global map of farm sizes at the subnational scale

(27). Since data for the economic size of a farm was not available, we used a Human Development Index (HDI) geospatial dataset of sub-nationally representative statistics from household surveys (145). The lowest 40% of HDI scores within each country were intersected with the smallest 40% of farm sizes to define smallholders.

Next, we overlaid the smallholder map with a global irrigation dataset (146) to examine the amount of irrigation used by smallholders compared to non-smallholders. We also overlaid it with a water scarcity dataset that was created by combining two datasets that represent regions experiencing at least one month of blue water scarcity (147) and less than 250 mm of annual rainfall (as per the IPCC 2007 definition for green water scarcity (148)). Our results show that smallholders have less irrigation coverage (39%) than non-smallholders (51%). This gap in irrigation narrows in water scarce regions, where smallholders had 27% of their cropland area irrigated and non-smallholders had 34% of their cropland irrigated.

While disparities between small-scale farmers and large-scale farmers are typically observed in local or national geographic contexts, our case study demonstrates that these inequalities can be monitored at the global level. Local analyses can target specific policy interventions, such as allocating public irrigation infastructure to certain locales; global monitoring of these disparities can assist in coordinated strategic planning for donors who work across regions (e.g., grant matching programs or pooled resources for targeted interventions). In addition, global monitoring of small-scale farmers' access to critical technologies can ensure govenments are held accountable to their SDG 2.3 commitments.

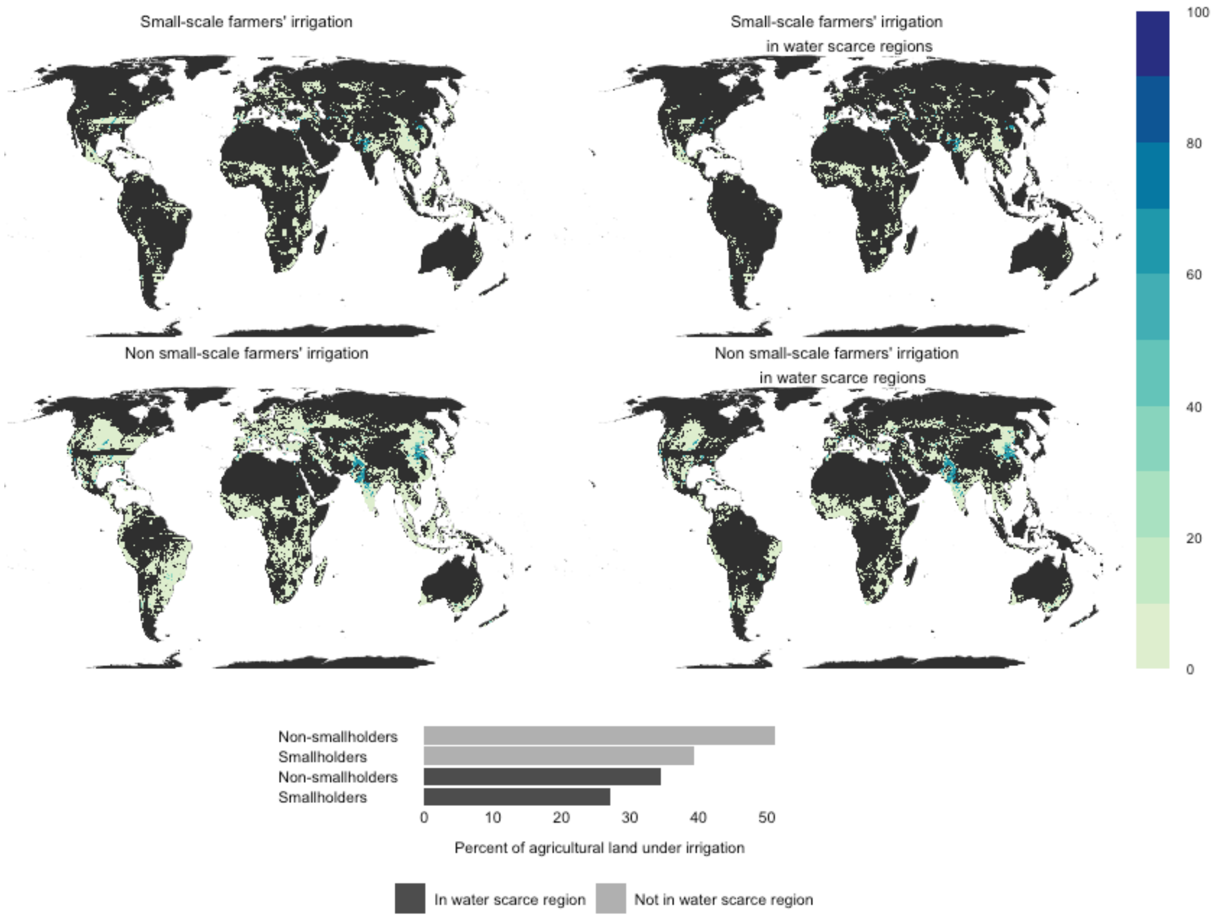

Figure 16: Maps show smallholders (top row) and non-smallholders' (bottom row) use of irrigation in all regions (left column) compared to water scarce regions (right column) at 10km resolution using Albers equal area projection. The percent of irrigation per pixel is given, where a dark blue pixel represents 100% of the crop area in the pixel is irrigated and a light green pixel represents 0% of the crop area in the pixel is irrigated. Smallholders are defined as the smallest 40% of farms in a country and with the lowest 40% human development index in a country (a proxy for agricultural revenue). National farm size data was downscaled using global field size maps. Water scarcity was defined as a pixel experiencing at least one month of blue water scarcity and less than 250 mm of rain per year.

5.5 Recommendations

We have four recommendations to improve the SDG 2.3's data efforts.

1. Existing and future data need to be made interoperable and geolocated. There are many household surveys and agricultural censuses that can be harmonized to provide baselines for small-scale farmers' productivity and incomes. Many of these surveys can be tied to spatial information at the finest resolution available (e.g., the survey sampling unit). There are efforts underway to further systematize future data collection on small-scale farms (e.g., the 50x2030 initiative and FAO's AGRIS). These efforts should include modules to georeference the field or household of the individual interviewed to allow interoperability with geospatial information. By linking social and environmental data, decision makers can allocate resources based on farmer specific contexts (e.g., ongoing crop monitoring efforts could transition from identifying areas with production losses to identifying populations experiencing losses). While any georeferencing will require additional resources and data management to avoid sensitive location information from being breached, there are examples in the public health community that have already established best practices for this exact type of data collection and dissemination (e.g., WHO DHS surveys).

2. The proposed FAO definition has empirical support for identifying impoverished farmers, but the definition still needs to be disaggregated by gender. While this definition can identify impoverished farmers, it may not capture gender, which is a central SDG 2.3 cross-cutting theme. We highlight that households with more female laborers had lower incomes; yet female labor was not correlated with FAO's dual definition of farm size and economic size.

3. The proposed binary FAO definition should be replaced with a multi-tiered definition. A multi-tiered definition would avoid misclassifications of farmers at the edge of a binary small-scale versus large farm dichotomy. We recommend classes for very-small, small, medium, large, and very-large farms. These multiple classes would enable more detailed monitoring of agrarian structure without additional data needed.

4. The terminology in SDG 2.3 target should be changed from "small-scale food producers" to "small-scale agricultural producers." The term "food" can be difficult to define for certain crops (e.g., many oil crops can be used for food and/or non-food purposes). In addition, only focusing on food may not include small-scale farmers that are producing non-food crops that are vital for their livelihoods (e.g., seed, animal feed, thatch, coir fiber for rope, etc.) The proposed FAO definition does not need to be changed to meet this rephrasing since it does not classify agricultural revenue or farm area based on food crops alone.

Chapter 6: Conclusion

This dissertation provides empirical support for understanding the contribution of small-scale farms to the global food system. Recently, there has been a wave of national and international policy calls to support small-scale farms, such as SDG 2.3 that aims to double their productivity and incomes (10, 11, 52). To ensure that international efforts are evidence-based, there is a need to understand the diversity and complexity of smallholder production systems as compared to non-smallholder systems. In this dissertation, I took a global perspective to empirically understand relationships between small-scale farmers, their contribution to the global food supply, and their socio-ecological impacts on outcomes and processes as compared to larger farms.

Global scale research on smallholders is relatively new, where the main body of empirical research began in 2015 (18–21, 149). This emerging research builds on a long history of case-studies and critical work on the socio-political inequities affecting smallholders' livelihoods (56, 150–152). Developments in technologies, advancements in research methods, and efforts to centralize and harmonize existing national sample surveys have made this recent global data synthesis work possible. In this dissertation, I have led and collaborated on efforts to harmonize new datasets, synthesized empirical literature from multiple disciplines, and combined social and environmental data. My aim was to add to the emerging global literature on small-scale farms.

The goal of this dissertation has been to understand the role of small-scale farms in the global food system by building upon previous localized and theoretical findings that underpin the current wave of international support for small-scale farmers. Chapters 2 through 4 provide empirical results on the role of small-scale farms in global crop and food production and their impact on several socio-ecological outcomes of farming systems. Chapter 5 places Chapter 4's results into the current policy context aimed to support small-scale farmers, and it presents future research directions. Table 7 presents my main findings.

We found that small-scale farms produce a large share of the world's food but much smaller than claimed by some authors. Farms under 2 ha in size, a conventional definition of small-scale farms, produce between 30-34% of the world's food on ~25% of the available agricultural land. In Chapter 2, we derived these statistics by harmonizing a novel dataset of crop production by farm size across 55 countries and over 150 crop species. Our finding is in line with two other studies that used geospatial data overlays of agricultural production and field sizes or average farm sizes (19, 20). The similarity of these three studies, which rely on different data and methods, provide strong evidence for small-scale farms production contributions to the global food supply.

Table 7: Dissertation's main findings.

Theme	Variable	Finding	Chapter
Production	Food produced	Farms under 2 ha produce ~1/3 of the global food supply, while farms under 5 ha produce nearly 1/2 the world's food and farms under 50 ha produce over 3/5 of the world's food.	2
	Non-food produced	Farms under 2 ha devote a greater proportion of their production to food, while farms over 1000 ha have up to the greatest proportion of post-harvest loss.	2
	Yield	Smaller farms have higher yields than larger farms, where an average 5% decrease in yields occurs per 1 ha increase in farm size. We find this relationship is largely moderated by differences in labor across farm sizes.	3
Environment	Crop diversity	Smaller farms account for greater crop diversity than larger farms, where farms under 2 ha account for ~40% of crop species richness in a given landscape.	2,3
	Non-crop diversity	Smaller farms promote greater non-crop biodiversity at the farm and landscape scales, where 77% of studies find smaller farms have more non-crop biodiversity than larger farms.	3
	Resource-use efficiency	We find no strong relationships between farm size and resource-use efficiency.	3
	Greenhouse gas emissions	We find a non-significant trend that smaller farms have lower GHG emission per unit of crop output than larger farms.	3
Socio-Economic	Economic productivity (profit per ha)	Smaller farms, on average, have higher economic productivity than larger farms, where an average 1% decrease in productivity occurs per 1 ha increase in farm size.	4
	Income (profit per capita)	Smaller farms, on average, have lower incomes than larger farms, where an average 2% increase in income occurs per 1 ha increase in farm size.	4
Policy	Small-scale farm definition	Actual and country relative farm size better explain farmer poverty compared to family labor and market orientation. A farm's country relevant economic size best relates to poverty of the definitions examined.	4
	Doubling incomes as an SDG target	We estimate that currently ~40% of farmers live above their national poverty lines. The SDG 2.3 target of doubling farmers' incomes would result in ~60% of farmers living above their national poverty lines. SDG 2.3 is not an aggressive enough target to transition farmers out of poverty.	4

Smaller farms devote a greater proportion of their production to food, while farms over 1000 ha have up to the greatest proportion of post-harvest loss. We found nearly 60% of small farms' (< 2 ha) production is allocated to food. A smaller percentage is allocated towards feed (12–16%), which was surprising since smallholders often engage in mixed crop-animal farming systems (45); this finding may be explained by the fact that smallholders likely rely more on rearing animals that graze on pasture compared to largeholders.

Smaller farms promote more crop and non-crop biodiversity than larger farms. From our dataset of 55 countries described in Chapter 2, we compared smaller and larger farms in the same landscape and found that smaller farms grew a greater number of crop species than larger farms (farms < 2 ha accounted for ~40% of crop species richness).

In Chapter 3, we conducted a meta-analysis of 30 studies (55 observations) and found a 77% probability that smaller farms were more likely to promote non-crop biodiversity at the farm and landscape level. This result held across different types of species and diversity metrics.

Our meta-analysis did not find a relationship between farm size and resource-use efficiency. Despite a widely acknowledged unequal access to credit, land, and markets for smallholders (73, 153), we found no difference between smaller and larger farm sizes and their resource-use efficiency across 29 studies (34 observations). Our results showed a non-significant trend that smaller farms had lower GHG emissions per unit output by crop than larger farms.

Smaller farms had higher productivity but lower incomes than larger farms. The higher productivity of smaller farms was consistent using two different metrics of productivity (i.e., yields and profit per hectare). In Chapter 3's meta-analysis of 18 studies (32 observations) we found for every 1 ha increase in farm size there was a 5% (95% CI: 1 to 9%) decrease in yield. In Chapter 4, we used harmonized data from 34 countries. We found that for every 1 ha increase in farm size there was a 1% (95% CI: 1.5 to 0.5%)[4] decrease in profit per hectare (in country relevant terms). We also found that for every 1 ha increase in farm size there was a 2% (95% CI: 1.6 to 2.3%) increase in income (in profit per capita terms). From our meta-analysis, we found that smaller farms are most likely more productive than larger farms due to the availability of unwaged family labor. Our results highlight that while smaller farms are more productive than larger farms, they have lower incomes than larger farms.

Country relevant farm size and economic size better identified impoverished farmers than alternative definitions tested. In Chapter 4, we used harmonized data from 30 countries to compare the relationship between different definitions of small-scale farms and poverty. We found that country relevant economic farm size explained the most variation in farmers' incomes (in profit per capita) compared to other tested small-scale farm definitions. Farm size (in hectare and country relevant terms) explained the second most amount of variation in farmers' incomes. The amount of family labor a farm relied on nor level of subsistence explained farm income levels. These results suggest that international initiatives can partially rely on the recently proposed dual definition of small-scale farms, which combines country relevant economic size and farm size, to

4 The same method that was used in Chapter 3's meta-analysis was applied to the statistics from Chapter 4 in order to convert them into relative changes. See Rodríguez-Barranco et al. 2017 for details.

identify impoverished farmers across geographic regions. While this definition can identify impoverished farmers, we found that it will not capture gender, which is a central SDG 2.3 cross-cutting theme in addition to a farmers' indigenous, religious, or caste status. Hence, any operational definition for small-scale farms will need to be disaggregated (see Chapter 5 for a more detailed policy discussion).

SDG 2.3's goal to double small-scale farmers' incomes is not an aggressive enough target to transition them out of poverty. In Chapter 4, we conducted simple scenarios to understand if the SDG 2.3 target of doubling small-scale farmers' incomes would transition farmers above their national poverty lines. In the 30 countries we analyzed, we found 38% of farming households currently live below their national poverty lines. If all farmers' incomes were doubled, regardless of the interventions used to double their incomes, only 60% of farmers would be living above their national poverty lines. Our simple scenarios revealed that while some farmers can transition out of poverty when doubling their incomes, a large majority of farmers will require greater support to make the transition.

References

1. Lowder SK, Skoet J, Raney T (2016) The Number, Size, and Distribution of Farms, Smallholder Farms, and Family Farms Worldwide. *World Dev* 87:16–29.

2. Altieri MA (2008) *Small farms as a planetary ecological asset: five key reasons why we should support the revitalisation of small farms in the global south*.

3. Chappell MJ, LaValle L a. (2009) Food security and biodiversity: can we have both? An agroecological analysis. *Agric Human Values* 28(1):3–26.

4. Perfecto I, Vandermeer J, Wright A (2009) *Nature's matrix: linking agriculture, conservation and food sovereignty* (Routledge).

5. FAO (2018) *The State of Food Security and Nutrition in the World* (Rome, Italy).

6. Wiggins S, Kirsten J, Llambi L (2010) The Future of Small Farms. *World Dev* 38(10):1341–1348.

7. Hazell P, Poulton C, Wiggins S, Dorward A (2010) The Future of Small Farms: Trajectories and Policy Priorities. *World Dev* 38(10):1349–1361.

8. Hazell PBRP (2005) Is there a future for small farms? *Agric Econ* 32:93–101.

9. Eastwood R, Lipton M, Newell A (2010) *Farm Size* (Elsevier B.V.). 1st Ed. doi:10.1016/S1574-0072(09)04065-1.

10. Sustainable Development Solutions Network (2015) *Indicators and a monitoring framework for Sustainable Development Goals - Launching a data revolution for the SDGs*.

11. Neufeld JL (2014) *The State of Food and Agriculture* doi:9789251073179.

12. Berdegué JA, Fuentealba R (2012) Latin America: The State of Smallholders in Agriculture. *Conference on New Directions for Smallholder Agriculture, IFAD Rome*, pp 1–40.

13. Feldmann C, Hamm U (2015) Consumers ' perceptions and preferences for local food : A review. *Food Qual Prefer* 40:152–164.

14. Meas T, Hu W, Batte MT, Woods TA, Ernst S (2015) Substitutes or complements? Consumer preference for local and organic food attributes. *Am J Agric Econ* 97(4):1044–1071.

15. Hu W, Batte MT, Woods T, Ernst S (2012) Consumer preferences for local production and other value-added label claims for a processed food product. 39(April 2018):489–510.

16. Dorottya S (2017) Determining the target groups of Hungarian short food supply chains based on consumer attitude and socio-demographic factors. *Agric Econ* 119:115–122.

17. Moon W, Pino G (2018) Do U.S. citizens support government intervention in agriculture? Implications for the political economy of agricultural protection. *Agric Econ (United Kingdom)* 49(1):119–129.

18. Graeub BE, et al. (2016) The State of Family Farms in the World. *World Dev* 87:1–15.

19. Samberg LH, Gerber JS, Ramankutty N, Herrero M, West PC (2016) Subnational distribution of average farm size and smallholder contributions to global food production. *Environ Res Lett* 11(12):124010.

20. Herrero M, et al. (2017) Farming and the geography of nutrient production for human use: a transdisciplinary analysis. *Lancet Planet Heal* 1(1):e33–e42.

21. Fritz S, et al. (2015) Mapping global cropland and field size. *Glob Chang Biol* 21(5):1980–1992.

22. Sibhatu KT, Krishna V V., Qaim M (2015) Production diversity and dietary diversity in smallholder farm households. *Proc Natl Acad Sci* 112(34):10657–10662.

23. Rudel T, et al. (2016) Do Smallholder, Mixed Crop-Livestock Livelihoods Encourage Sustainable Agricultural Practices? A Meta-Analysis. *Land* 5:6.

24. Rosset PM, Martínez-Torres ME, Martinez-Torres ME (2012) Rural social movements and agroecology: Context, theory, and process. *Ecol Soc* 17(3). doi:10.5751/ES-05000-170317.

25. Wittman H (2011) Food Sovereignty: A New Rights Framework for Food and Nature? *Environ Soc Adv Res* 2(1):87–105.

26. HLPE (2013) Investing in Smallholder Agriculture for Food Security. A Report by the High Level Panel of Experts on Food Security and Nutrition of the Committee on World Food Security. (June):111.

27. Mehrabi, Z., Ricciardi, V., McDowell M A global baseline for data-driven agronomy interventions.

28. Khalil CA, Conforti P, Ergin I, Gennari P (2017) Defining Small-Scale Food Producers To Monitor Target 2.3. of the 2030 Agenda for Sustainable Development. *FAO Stat Work Pap Ser* 17(12). Available at: www.fao.org/3/a-i6858e.pdf.

29. Ricciardi V, Ramankutty N, Mehrabi Z, Jarvis L, Chookolingo B (2018) How much of our world's food do smallholders produce? *Glob Food Sec* 17:64–72.

30. ETC (2009) Who will feed us? Questions for the food and climate crisis. *Commun Issue 102* (November):34.

31. Maass Wolfenson KD (2013) Coping with the food and agriculture challenge : smallholders ' agenda. *2012 United Nations Conf Sustain Dev* (April):47.

32. Badstue LB, Bellon MR, Berthaud J, Ramírez A (2005) Collective Action for the Conservation of On-Farm Genetic Diversity in a Center of Crop Diversity : An Assessment of the Role of Traditional Farmers ' Networks.

33. Conway G (2011) On Being a Smallholder. *Conf new Dir Smallhold Agric IFAD Rome*:1–17.

34. Horrigan L, Lawrence RS, Walker P (2002) How sustainable agriculture can address the environmental and human health harms of industrial agriculture. *Environ Health Perspect* 110(5):445–456.

35. Naylor R, et al. (2005) Losing the Links Between. *Science (80-)* (December):1621–1622.

36. IFAD, UNEP (2013) Smallholders, food security and the environment. *Rome Int Fund Agric Dev*:54.

37. UN (2015) Sustainable Development Goals (SDGs). Available at: https://sustainabledevelopment.un.org/topics/sustainabledevelopmentgoals# [Accessed January 1, 2017].

38. Hodges RJJ, Buzby JCC, Bennet B, Bennett B (2011) Postharvest losses and waste in developed and less developed countries: Opportunities to improve resource use. *J Agric Sci* 149(S1):37–45.

39. Tefera T (2012) Post-harvest losses in African maize in the face of increasing food shortage. *Food Secur* 4(2):267–277.

40. Rahman S, Kazal MMH (2015) Determinants of crop diversity in the regions of Bangladesh (1990-2008). *Singap J Trop Geogr* 36(1):83–97.

41. Van den Berg MM, et al. (2007) The impact of increasing farm size and mechanization on rural income and rice production in Zhejiang province, China. *Agric Syst* 94(3):841–850.

42. FAOSTAT (2017) Food and Agriculture Organization of the United Nations FAOSTAT Database.

43. Cassidy ES, West PC, Gerber JS, Foley JA (2013) Redefining agricultural yields: from tonnes to people nourished per hectare. *Environ Res Lett* 8(3):034015.

44. Manly BFJ (2006) *Randomization bootstrap and monte carlo methods in biology, Third Edition* (Chapman and Hall/CRC).

45. Smith J, et al. (2012) Beyond milk, meat, and eggs: Role of livestock in food and nutrition security. *Anim Front* 3(1):6–13.

46. Kummu M, et al. (2012) Lost food, wasted resources: Global food supply chain losses and their impacts on freshwater, cropland, and fertiliser use. *Sci Total Environ* 438:477–489.

47. Cohn AS, et al. (2017) Smallholder Agriculture and Climate Change. *Annu Rev Environ Resour* 42(1):annurev-environ-102016-060946.

48. Assunção JJ, Braido LHB (2007) Testing household-specific explanations for the inverse productivity relationship. *Am J Agric Econ* 89(4):980–990.

49. Fifanou VG, Ousmane C, Gauthier B, Brice S (2011) Traditional agroforestry systems and biodiversity conservation in Benin (West Africa). *Agrofor Syst* 82(1):1–13.

50. Keleman A, Hellin J, Flores D (2013) Diverse Varieties and Diverse Markets: Scale-related Maize "Profitability Crossover" in the Central Mexican Highlands. *Hum Ecol* 41(5):683–705.

51. Baiphethi MN, Jacobs PT (2009) The contribution of subsistence farming to food security in South Africa. *Agrekon* 48(4):459–482.

52. UNFCCC Intended Nationally Determined Contributions (INDCs). *2015*. Available at: https://www4.unfccc.int/sites/submissions/indc/Submission Pages/submissions.aspx [Accessed February 6, 2019].

53. Konvicka M, Benes J, Polakova S (2016) Smaller fields support more butterflies: comparing two neighbouring European countries with different socioeconomic heritage. *J Insect Conserv* 20(6):1113–1118.

54. Haji J (2007) Production efficiency of smallholders' vegetable-dominated mixed farming system in eastern Ethiopia: A non-parametric approach. *J Afr Econ* 16(1):1–27.

55. Barrett CB, Bellemare MF, Hou JY (2010) Reconsidering Conventional Explanations of the Inverse Productivity-Size Relationship. *World Dev* 38(1):88–97.

56. Sen AK (1962) An Aspect of Indian Agriculture. *Econ Wkly* 14(4–6):243–246.

57. Chayanov AV V. (1926) *The Theory of Peasant Cooperatives* (Ohio State University Press, Columbus, OH).

58. Otsuka K, Liu Y, Yamauchi F (2016) Growing advantage of large farms in Asia and its

implications for global food security. *Glob Food Sec* 11:5–10.

59. Rada NE, Fuglie KO (2019) New perspectives on farm size and productivity. *Food Policy* 84. doi:10.1016/j.foodpol.2018.03.015.

60. Smith RK, Jennings NV (2005) A quantitative analysis of the abundance and demography of European hares. 35(1):1–24.

61. Moher D, Liberati A, Tetzlaff J, Altman DG (2009) Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *Bmj* 339(b2535). doi:10.1136/bmj.b2535.

62. Batte MT, Ehsani MR (2006) The economics of precision guidance with auto-boom control for farmer-owned agricultural sprayers. *Comput Electron Agric* 53(1):28–44.

63. Ouin A, Burel F (2002) Influence of herbaceous elements on butterfly diversity in hedgerow agricultural landscapes. *Agric Ecosyst Environ* 93(1–3):45–53.

64. Bravo-Monroy L, Tzanopoulos J, Potts SGG (2015) Ecological and social drivers of coffee pollination in Santander, Colombia. *Agric Ecosyst Environ* 211:145–154.

65. Brown PW, Schulte LA (2011) Agricultural landscape change (1937-2002) in three townships in Iowa, USA. *Landsc Urban Plan* 100(3):202–212.

66. Teshome A, Patterson D, Asfaw Z, Dalle S, Torrance JK (2016) Changes of Sorghum bicolor landrace diversity and farmers' selection criteria over space and time, Ethiopia. *Genet Resour Crop Evol* 63(1):55–77.

67. Gedebo A, Appelgren M, Bjornstad A, Tsegaye A (2007) Analysis of indigenous production methods and farm-based biodiversity of amochi (Arisaema schimperian, Schott) in two sub-zones of Southern Ethiopia. *Genet Resour Crop Evol* 54(7):1429–1436.

68. Altman DG, et al. (1998) Predictors of crop diversification: a survey of tobacco farmers in North Carolina (USA). *Tob Control* 7(4):376–382.

69. Külekçi M (2010) Technical efficiency analysis for oilseed sunflower farms: A case study in Erzurum, Turkey. *J Sci Food Agric* 90(9):1508–1512.

70. Latruffe L, Balcombe K, Davidova S, Zawalinska K (2005) Technical and scale efficiency of crop and livestock farms in Poland: Does specialization matter? *Agric Econ* 32(3):281–296.

71. Ullah A, Perret SR (2014) Technical- and environmental-efficiency analysis of irrigated cotton-cropping systems in Punjab, Pakistan using data envelopment analysis. *Environ Manage* 54(2):288–300.

72. Binici T, Zulauf CR, Kacira OO, Karli B (2006) Assessing the efficiency of cotton production on the Harran Plain, Turkey. *Outlook Agric* 35(3):227–232.

73. Deininger K, Zegarra E, Lavadenz I (2003) Determinants and impacts of rural land market activity: Evidence from Nicaragua. *World Dev* 31(8):1385–1404.

74. Deininger K, Byerlee D (2012) The rise of large farms in land abundant countries: Do they have a future? *World Dev* 40(4):701–714.

75. Stifel D, Minten B (2008) Isolation and agricultural productivity. *Agric Econ* 39(1):1–15.

76. Rada N, Wang C, Qin L (2015) Subsidy or market reform? Rethinking China's farm consolidation strategy. *Food Policy* 57:93–103.

77. Barrett CB, Bellemare MF, Hou JY (2010) Reconsidering Conventional Explanations of the Inverse Productivity Size Relationship. *World Dev* 38(1):88–97.

78. Carter BMR (1984) Identification of the Inverse Relationship between Farm Size and Productivity : An Empirical Analysis of Peasant Agricultural Production Author (s): Michael R . Carter Source : Oxford Economic Papers , New Series , Vol . 36 , No . 1 (Mar ., 1984), pp. 36(1):131–145.

79. Henderson H (2015) Considering Technical and Allocative Efficiency in the Inverse Farm Size-Productivity Relationship. *J Agric Econ* 66(2):442–469.

80. Chen Z, Huffman WE, Rozelle S (2011) Inverse relationship between productivity and farm size: The case of China. *Contemp Econ Policy* 29(4):580–592.

81. Henderson H (2014) Structural transformation and smallholder agriculture: An information-theoretic analysis of the Nicaraguan case. *Agric Econ (United Kingdom)* 45(4):443–458.

82. Fan S, Chan-Kang C (2005) Is small beautiful? Farm size, productivity, and poverty in Asian agriculture. *Agric Econ* 32(SUPPL. 1):135–146.

83. Sen AK (1964) Size of Holdings and Productivity. *Econ Wkly* (February):323–326.

84. Tukaew S, Datta A, Shivakoti GP, Jourdain D (2016) Production Practices Influenced Yield and Commercial Cane Sugar Level of Contract Sugarcane Farmers in Thailand. *Sugar Tech* 18(3):299–308.

85. Newell A, Pandya K, Symons J (1997) Farm Size and the Intensity of Land Use in Gujarat. *Oxf Econ Pap* 49(2):307–315.

86. Verschelde M, D'Haese M, Rayp G, Vandamme E (2013) Challenging small-scale farming: A non-parametric analysis of the (inverse) relationship between farm productivity and farm

size in burundi. *J Agric Econ* 64(2):319–342.

87. Carletto C, Savastano S, Zezza A (2013) Fact or artifact: The impact of measurement errors on the farm size–productivity relationship. *J Dev Econ* 103:254–261.

88. Abay KA., Abate GT., Barrett CB., Tanguy B (2018) Correlated non-classical measurement errors, 'second best' policy inference and the inverse size-productivity relationship in agriculture. IFPRI Disc.

89. Lesiv M, et al. (2019) Estimating the global distribution of field size using crowdsourcing. (September 2018):174–186.

90. Graesser J, Ramankutty N (2017) Remote Sensing of Environment Detection of cropland field parcels from Landsat imagery. *Remote Sens Environ* 201:165–180.

91. Jonsen ID, Fahrig L (1997) Response of generalist and specialist insect herbivores to landscape spatial structure. *Landsc Ecol* 12(3):185–197.

92. Ahrenfeldt EJJ, et al. (2015) Pollinator communities in strawberry crops – variation at multiple spatial scales. *Bull Entomol Res* 105(04):497–506.

93. Concepción ED, Fernandez-González F, Díaz M (2012) Plant diversity partitioning in Mediterranean croplands: Effects of farming intensity, field edge, and landscape context. *Ecol Appl* 22(3):972–981.

94. Belfrage K, Björklund J, Salomonsson L (2005) The effects of farm size and organic farming on diversity of birds, pollinators, and plants in a Swedish landscape. *Ambio* 34(8):582–588.

95. Horgan FG (2009) Invasion and retreat: Shifting assemblages of dung beetles amidst changing agricultural landscapes in central Peru. *Biodivers Conserv* 18(13):3519–3541.

96. Schai-Braun SC, Hacklander K (2014) Home range use by the European hare (Lepus europaeus) in a structurally diverse agricultural landscape analysed at a fine temporal scale. *Acta Theriol (Warsz)* 59(2):277–287.

97. Lovell ST, Mendez VE, Erickson DL, Nathan C, DeSantis S (2010) Extent, pattern, and multifunctionality of treed habitats on farms in Vermont, USA. *Agrofor Syst* 80(2):153–171.

98. Pekin BK (2016) Anthropogenic and topographic correlates of natural vegetation cover within agricultural landscape mosaics in Turkey. *Land use policy* 54:313–320.

99. Hansen ZK, Libecap GD, Hansen ZK, Libecap GD (2004) Small Farms, Externalities, and the Dust Bowl of the 1930s. *J Polit Econ* 112(3):665–694.

100. Sanza M a., Traba J, Morales MB, Rivera D, Delgado MP (2012) Effects of landscape, conspecifics and heterospecifics on habitat selection by breeding farmland birds: the case of the Calandra Lark (Melanocorypha calandra) and Corn Bunting (Emberiza calandra). *J Ornithol* 153(2):525–533.

101. Panek M (2016) The Effect of Agricultural Landscape Structure on Food Resources and Survival of Grey Partridge Perdix Perdix Chicks in Poland Author (s): Marek Panek Source : Journal of Applied Ecology , Vol . 34 , No . 3 (Jun ., 1997), pp . 787-792 Published by : B. 34(3):787–792.

102. Rosin ZM, et al. (2012) Landscape structure, human disturbance and crop management affect foraging ground selection by migrating geese. *J Ornithol* 153(3):747–759.

103. Zimmerer KS (2011) Geographies of Seed Networks for Food Plants (Potato , Ulluco) and Approaches to Agrobiodiversity Conservation in the Andean Countries. *Soc Nat Resour An Int J* 16:583–601.

104. Bicksler A, et al. (2012) Methodologies for strengthening informal indigenous vegetable seed systems in northern Thailand and Cambodia. *Acta Hortic* 958:67–74.

105. Pautasso M, et al. (2012) Seed exchange networks for agrobiodiversity conservation. A review. *Agron Sustain Dev* 33(1):151–175.

106. Coomes OT, et al. (2015) Farmer seed networks make a limited contribution to agriculture? Four common misconceptions. *Food Policy* 56:41–50.

107. McCord PF, Cox M, Schmitt-Harsh M, Evans T (2015) Crop diversification as a smallholder livelihood strategy within semi-arid agricultural systems near Mount Kenya. *Land use policy* 42:738–750.

108. Clark M, Tilman D (2017) Comparative analysis of environmental impacts of agricultural production systems, agricultural input efficiency, and food choice. *Environ Res Lett* 12(6):064016.

109. Manjunatha A V., Anik AR, Speelman S, Nuppenau EA (2013) Impact of land fragmentation, farm size, land ownership and crop diversity on profit and efficiency of irrigated farms in India. *Land use policy* 31:397–405.

110. Alene AD, Hassan RM (2003) The determinants of farm-level technical efficiency among adopters of improved maize production technology in western Ethiopia. *Agrekon* 42(1):1–14.

111. Gaurav S, Mishra S (2015) Farm Size and Returns to Cultivation in India: Revisiting an Old Debate. *Oxford Dev Stud* 43(2):165–193.

112. Frelat R, et al. (2015) Drivers of household food availability in sub-Saharan Africa based on big data from small farms. *Proc Natl Acad Sci U S A* 113(2):1518384112-.

113. Josephson AL, Ricker-Gilbert J, Florax RJGMGM (2014) How does population density influence agricultural intensification and productivity? Evidence from Ethiopia. *Food Policy* 48:142–152.

114. Chand R, Prasanna P a L, Singh A (2011) Farm size and productivity: Understanding the strengths of smallholders and improving their livelihoods. *Econ Polit Wkly* XLVI(26/27):5–11.

115. Barthel S, Crumley C, Svedin U (2013) Bio-cultural refugia-Safeguarding diversity of practices for food security and biodiversity. *Glob Environ Chang* 23(5):1142–1152.

116. Kremen C, Iles A, Bacon C (2012) Diversified Farming Systems : An Agroecological , Systems-based. *Ecol Soc* 17(4).

117. Seufert V, Ramankutty N (2017) Many shades of gray - The context-dependent performance of organic agriculture. *Sci Adv* 3(3):e1602638.

118. Harrison JL, Getz C (2015) Farm size and job quality: mixed-methods studies of hired farm work in California and Wisconsin. *Agric Human Values* 32(4):617–634.

119. Agresti A (2002) *Categorical Data Analysis* (Wiley). 3rd Editio.

120. Christensen RHB (2015) Analysis of ordinal data with cumulative link models - estimation with the R -package ordinal. 1–31.

121. Becker BJ, Wu M-J (2007) The Synthesis of Regression Slopes in Meta-Analysis. *Stat Sci* 22(3):414–429.

122. Viechtbauer W (2015) Conducting Meta-Analyses in R with the metafor Package . *J Stat Softw* 36(3). doi:10.18637/jss.v036.i03.

123. Rodríguez-Barranco M, Tobías A, Redondo D, Molina-Portillo E, Sánchez MJ (2017) Standardizing effect size from linear regression models with log-transformed variables for meta-analysis. *BMC Med Res Methodol* 17(1):1–9.

124. Mahon N, Crute I, Simmons E, Islam MM (2017) Sustainable intensification – "oxymoron" or "third-way"? A systematic review. *Ecol Indic* 74:73–97.

125. Harris D (2019) Intensification benefit index: How much can rural households benefit from agricultural intensification? *Exp Agric* 55(2):273–287.

126. Rasmussen LV, et al. (2018) Social-ecological outcomes of agricultural intensification. *Nat*

Sustain 1(6):275–282.

127. Ricciardi V, Mehrabi Z, Wittman H, James D, Ramankutty N (2019) Smaller farms are higher yielding and more biodiverse than larger farms: A systematic review and meta-analysis. *Forthcoming*.

128. Abler DG, Sukhatme VA (2006) The "Efficient but Poor" Hypothesis. *Rev Agric Econ* 28(3):338–343.

129. Dahal BM, Nyborg I, Sitaula BK, Bajracharya RM (2009) Agricultural intensification: Food insecurity to income security in a mid-hill watershed of Nepal. *Int J Agric Sustain* 7(4):249–260.

130. Dawson N, Martin A, Sikor T (2016) Green Revolution in Sub-Saharan Africa: Implications of Imposed Innovation for the Wellbeing of Rural Smallholders. *World Dev* 78:204–218.

131. Barrett CB (1996) On price risk and the inverse farm size-productivity relationship. *J Dev Econ* 51(2):193–215.

132. van Vliet J a., et al. (2015) De-mystifying family farming: Features, diversity and trends across the globe. *Glob Food Sec* 5:11–18.

133. Garner E, de la O Campos AP (2014) Identifying the family farm. *ESA Work Pap* (14):1–30.

134. WorldBank, FAO, IFAD (2009) *Gender in agriculture sourcebook* (Washington, DC).

135. Reddy S, Lahoti R (2016) $1.90 a Day: What Does it Day? *New Left Rev* 97:106–127.

136. Pinheiro JC, Bates DM (2000) *Mixed effects models in S and S-Plus* (Springer, New York, NY). 5th Ed.

137. Faraway JJ (2006) Extending the linear model with R: generalized linear, mixed effects and nonparametric regression models. 1–28.

138. Maine N, Lowenberg-DeBoer J, Nell WT, Alemu ZG (2010) Impact of variable-rate application of nitrogen on yield and profit: A case study from South Africa. *Precis Agric* 11(5):448–463.

139. Lastarria-cornhiel S (2006) Feminization of agriculture: Trends and driving forces. *Backgr Pap world Dev Rep 2008*:1–22.

140. Collins JL (2000) Tracing social relations in commodity chains: the case of grapes in Brazil. *Commodities and Globalization: Anthropological Perspectives.*

141. Barrientos S, Dolan C, Tallontire A (2003) A gendered value chain approach to codes of conduct in African horticulture. *World Dev* 31(9):1511–1526.

142. Harris D, Orr A (2014) Is rainfed agriculture really a pathway from poverty? *Agric Syst* 123:84–96.

143. Lowder SK, Skoet J, Singh S (2014) What do we really know about the number and distribution of farms and family farms in the world ? *ESA Work Pap* 14–02(14):38.

144. USAID (2018) Famine Early Warning Systems (FEWS) Network. Available at: http://fews.net/ [Accessed October 28, 2018].

145. Smits J, Permanyer I (2019) Data descriptor: The subnational human development database. *Sci Data* 6:1–15.

146. Salmon JM, Friedl MA, Frolking S, Wisser D, Douglas EM (2015) Global rain-fed, irrigated, and paddy croplands: A new high resolution map derived from remote sensing, crop inventories and climate data. *Int J Appl Earth Obs Geoinf* 38:321–334.

147. Hoekstra AY, Mekonnen MM, Chapagain AK, Mathews RE, Richter BD (2012) Global monthly water scarcity: Blue water footprints versus blue water availability. *PLoS One* 7(2). doi:10.1371/journal.pone.0032688.

148. IPCC (2007) *Climate Change 2007: impacts, adaptation and vulnerability. Contribution of Working Group II to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change.* eds Parry OFC, Palutikof JP, Hanson C., van der Linden PJ.

149. Lowder SK, Carisma B, Skoet J (2015) Who Invests How Much in Agriculture in Low-and Middle-Income Countries? An Empirical Review. *Eur J Dev Res* 27:371–390.

150. Netting R (1974) Agrarian ecology. *Annu Rev Anthropol* 3(1):21–56.

151. Boserup E (1965) *The Conditions of Agricultural Growth: The Economics of Agrarian Change Under Population Pressure* (Routledge, London).

152. Weiler AM, et al. (2015) Food sovereignty, food security and health equity: A meta-narrative mapping exercise. *Health Policy Plan* 30(8):1078–1092.

153. Arene CJ (1992) Comparative Economics of Maize and Rice Production among Resource-Poor Farmers in Anambra State of Nigeria. *African Dev Rev* 4(1):102–113.

154. Dorsey B (1999) Agricultural Intensification, Diversification, and Commercial Production among Smallholder Coffee Growers in Central Kenya*. *Econ Geogr* 75(2):178–195.

155. Bevis LEM, Barrett CB (2016) Close to the Edge: Do Behavioral Explanations Account for

the Inverse Productivity Relationship?

156. Kermath BM, Bennett BC, Pulsipher LM (2014) Food Plants in the Americas: A Survey of the Domesticated, Cultivated, and Wild Plants Used for Human Food in North, Central and South America and the Caribbean. *Unpubl manuscript, Univ Wisconsin Oshkosh, Oshkosh*.

157. Fern K, Fern A, Morris R (2018) Useful Tropical Plants Database. Available at: http://tropical.theferns.info/ [Accessed August 30, 2017].

Appendices

Appendix A Chapter 2 supplemental information

A.1. Previous studies estimating global crop and food production by farm size

We discuss four previous studies that quantified either the number and area of farms under smallholder management (i.e., Lowder et al. 2016; Graeub et al. 2016) or how food production varies by farm size (i.e., Herrero et al. 2017; Samberg et al. 2016) [Lowder, Graeub, Herrero, and Samberg hereafter, respectively]. Due to the difficulty in compiling global statistics on agriculture, each of these studies faced issues with available data or used modelling assumptions that affected their final estimates. This supplemental overview discusses these issues in order to both differentiate our study as well to have a central location in the literature that compares each study.

Lowder was the first study to compile global data on the percentage of smallholders in the world in terms of how many farms are in each farm size class, how much land each farm size class manages, and what percentage of farms are family farms. This dataset was based on the World Census of Agriculture (WCA) farm size distributions, which are country reported national estimates. Yet, many nationally reported agricultural datasets may have varying reliability and include inconsistent years of data with some countries' most recent statistics dating back to the 1960s (average year: 1997). Lowder also included distributions of total farm area rather than cropping area by farm size; there is no actual way to distinguish how much area was cropped or harvested when using this dataset.

Graeub quantified the number of family farms in the world and their global production contributions. While many small farms are family farms, the authors' data and examples point out that family farms and farm size should not be conflated. For example, family farms in Graeub's case study of Brazil may be family owned but are large in size (while ~85% of farms in Brazil are family owned covering ~25% of agricultural land, only 21% of farms are less than 2 ha in size and cover only 0.25% of the agricultural area). Graeub estimated ~98% of all farms globally are family farms, collectively managing 53% of all cropland, and meeting an estimated 36–114% of domestic caloric requirements. There were discrepancies between Lowder and Graeub family farm estimates: according to Lowder, family farms made up an average of 73.5% of agricultural land in the 77 countries they analyzed, while Graeub found an average of 53% in 105 countries they analyzed. This discrepancy may be due to the more nuanced, per country definition of family farms presented by Graeub. Lowder did not offer a definition for family farms in their methodology, but relied on government reports.

Samberg was the first study to estimate the global food production contributions of smallholders in an analysis of 83 countries in smallholder dominant areas of Latin America, sub-Saharan Africa, and South and East Asia. They estimated that farms under 5 ha managed 30% of global agricultural land and produced ~53% of global food calories in their sample. This study was a valuable initial step in understanding how much food smallholders produced, but may have had inaccuracies due to the following methodological reasons. First, their study relied on an estimate of mean size in a given administrative unit (referred to as 'mean agricultural area' or MAA in their article), and therefore did not capture size distributions within; using the mean may bias results because farm size distributions are highly skewed (Lowder). Second, their metric MAA included the area of cropland and pasture, and not the entire farm, so they may be overestimating the area of farms under smallholder management as larger farms are more likely to have non-agricultural areas (e.g., fallow area, water sources, infrastructure, etc.); we ran into a similar issue for ~5% of our dataset, where we had to assume that the size of a farm's aggregated plots equated to its farm size (we test for this assumption in Appendix A). Third, they did not account for the fact that the types of crops grown potentially vary by farm size within administrative units and hence variations in production for food versus other uses. Their study simply used the geographic relationships between crop types and MAA for administrative units. Several agricultural economic and agronomic studies have found that different farm sizes may grow different species and varieties of crops due to market and input access, as well as the subsistence focus of many smallholders (40, 48, 154). Fourth, their data may over estimate smallholder production since their 'global' sample only contains low-income, smallholder dominant countries: their sample of 83 countries accounts for 90% of all farmers, but only represents 35% of global cropland. Fifth, the authors attempt to disaggregate national crop production to their MAAs evenly; meaning, they assumed every farm size within an administrative unit has the same yields. There is a long history in the agricultural economic literature that suggests smaller farms have higher yields; a phenomenon coined as the 'inverse farm size-yield relationship' (79, 83, 155). In our paper, we also had to rely on constant yields for ~60% of our data. In the following sections, we show tests for this bias and determine it may not be an issue at this scale of analysis. Finally, their analysis omitted non-family farms by relying on household surveys to estimate where smallholders live, resulting in another potential source of overestimation since their data came from family-farms, which are often associated with smallholders (Graeub). We faced a similar issue for 22.5% of crop production in our dataset; see the following sections for detailed discussion on our dataset and the family farm bias. We were not able to provide bias estimates due to the conflicting findings between Graeub and Lowder on how much cropland is under family farm management as detailed above.

The data and methods used in Herrero's analysis also needs to be further verified due to potential issues with two of their underlying datasets. The first is Lowder farm size distributions, as detailed above. Herrero attempted to update the farm size distribution dataset where possible, but relied on imputing missing data by using regional farm size distributions. Neighboring countries may not be the best indicator of farm size since national land reform policies, colonial land grab legacies, market integration, and subsidy support are main drivers of farm size rather than pure geographic proximity (9). The second dataset (also used by Samberg) that may have introduced error was a crowd sourced field size dataset heavily reliant on interpolation (the data product was 99.99% interpolated[5]) and based on a qualitative scale of five farm sizes from very small to large that did not indicate actual hectares per category (21). In terms of methodology, Herrero made an explicit and transparent assumption about the relationship between field and farm size, which may have led to over estimates of the number of farms and cropping area under smallholder management. They used Lowder's national farm size distributions to estimate sub-national level farm size distributions, then matched these distributions with the crowd sourced field sizes; however, the

5 Fritz et al. (2015) used 13,963 samples of 1 km2 pixels of the global 1,554,216,620 km2 cropland (FAOSTAT 2005 (http://www.fao.org/faostat/en/) estimate of arable land and permanent crops as similarly defined by Fritz et al.), hence, their dataset was based on a 0.0009% sample of total global cropland (100* 13,963 / 1,554,216,620).

manner in which they used these distributions was not transparent in the available written methodology. Their assumption that national level distributions of farm size can be downscaled may have introduced error in their attribution of certain fields to farms. More importantly, assigning the national level farm size relationships to determine a field to farm size functional relationship may make sense in the case of large fields, which can only belong to larger farms. Yet, many larger farms may either hold multiple, non-contiguous smaller fields or they may just cultivate fewer large fields. Assigning small fields to small farms may under/over-estimate the amount of land under smallholder cultivation. Finally, Herrero also relied on the same constant yield bias that Samberg and our study did, where they applied per pixel level yields evenly to each farm sizes' cropped area to link the cropland datasets to crop production. Regardless, Herrero was the first to provide a complete global estimate of crop production by farm size and to illustrate an important link between smallholders and micro-nutrient production.

Critically, each prior attempt to quantify the amount of global food and crop production did not incorporate a validation method to assess whether they were assigning an agricultural field to the appropriately sized farm since at the time of their analysis there was no globally representative dataset that describes this relationship. Our dataset, while not globally contiguous can provide this needed validation. Additionally, we also ran into similar methodological issues due to data constraints, such as a partial reliance on constant yields (representing ~60% of our data), omitting non-family farms (22.5% of our data), and using aggregated plot size as a proxy for farm size (~5% of our data). In the following sections, we provide sensitivity tests and detailed explanation of each of these biases to be transparent about our dataset's limitations, offer insight into past attempts relying on these same biases, and offer guidance for any future work. Lastly, all previous attempts to quantify smallholder production, including our study, has defined smallholders by a set farm size class for the entire world (e.g., < 2 ha). This definition may also be problematic because it does not account for alternate definitions of farm size, such as economic output or relative farm size per country. While our dataset cannot be used to recalculate farm size in economic terms, it could be used for a reanalysis of farm size in country relevant terms.

A.2. Dataset construction

This dataset was built to provide uncertainty estimates for the percentage of food produced by farms of different sizes globally. We constructed this dataset by harmonizing agricultural censuses and nationally representative household sample surveys that directly measured crop production and/or cropping area[6] by farm size. This dataset is a convenience sample of 55 countries with 45 countries having sub-national resolution.

Our dataset captures ~51.1% of global crop production and ~52.9% of global cropland area (i.e., arable land and permanent crop area as reported in the Food and Agricultural Organization's statistical database [FAOSTAT hereafter](42)). The primary sources are agricultural census data (i.e., the majority of which used exhaustive sampling of the farming population, but not all response rates were 100%) or nationally representative sample surveys (i.e., with randomly stratified sampling of households in a country). These data were available at either the aggregated

6 Where there was no crop production by farm size data available we extracted farm size by either harvest area, cultivated area, crop area, or planted area. We will refer to this as 'cropping area' in this article.

level by administrative unit (34 countries) or at the non-aggregated, microdata level where data is available as anonymized individual household level records (21 countries, of which 18 were sample surveys and 3 were complete agricultural censuses) Figure S1). We document the source information, detail the methods for building this dataset, and describe its characteristics in this article to enable its use by the research community.

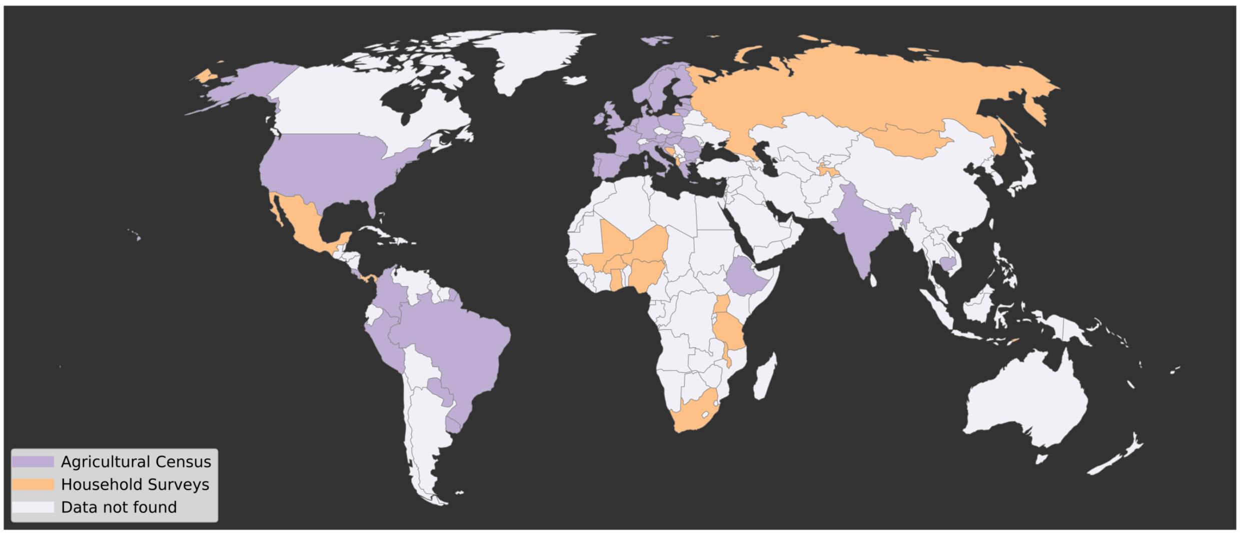

Figure S1: Map showing source of data derived from agricultural censuses (purple) or household surveys (orange) at the country level.

This database was harmonized across countries, 154 crop species, and farm size categories. Crop species and country names were matched with FAOSTAT by year to integrate with its extensive variable lists. The median year of the source data was from 2013, with the oldest source dataset from 2001 and the newest from 2015; each administrative unit contains data for the most recently available time point. We harmonized the farm size categories to match the World Census of

Agriculture (WCA) farm size categories: 0 to 1 ha, 1 to 2 ha, 2 to 5 ha, 5 to 10 ha, 10 to 20 ha, 20 to 50 ha, 50 to 100 ha, 100 to 200 ha, 200 to 500 ha, 500 to 1000 ha, and above 1000 ha.

We ran into several methodological issues when harmonizing the underlying the data needed to construct this dataset. In this article, we outline the assumptions made, and test the bias of these assumptions, such as applying constant yields across farm size classes to estimate production when only cropping area was available (representing ~60% of our data), omitting non-family farms when relying on household sample surveys (22.5% of our data), using aggregated plot size as a proxy for farm size (~5% of our data), and omitting crop species that were not able to be harmonized across countries or with the FAOSTAT crop species list.

In this supplemental material, we also provide details on the data collection and inclusion process, summary statistics, spatial coverage, and provide sensitivity tests and/or detailed explanations of each of the data harmonization assumptions we made. Our goal is to be transparent about our dataset's limitations, offer insight for other data harmonization projects relying on these same biases, and offer guidance for people wishing to use this data in their own work.

A.2.1. Methods for data selection

Inclusion criteria

We prescribed four inclusion criteria for this project. First, datasets needed to contain variables for farm size (where farm size was not available we relied on aggregate field size), and production per crop or cropping area per crop. Second, datasets needed to be nationally representative. Agricultural censuses or household sample surveys were used only when their sampling methodology was transparent and/or these datasets were used by the country's government for official statistics. We required the household surveys' sampling designs to be transparent, randomized at the appropriate administrative unit, and to provide sampling weights and expansion factors with details on their creation and intended application. Third, national numbers calculated from these datasets needed to be comparable with official national statistics. For many agricultural censuses, the sampling design and response rates were not available. Fourth, we only focused on surveys which included disaggregated data on crop species so that they could be matched according to FAOSTAT crop names and item codes. No aggregate category was used (e.g., 'roots and tubers' or 'fruit and vegetables').

We systematically searched several locations for agricultural datasets to compile our dataset. These sources included the World Bank microdata archives, EarthStat metadata, Living Standards Measurement Study (LSMS) surveys, and the Accelerated Data Program (see Table S1). We conducted our search on a per country basis either through each data archive's search capabilities where available, detailed search of each data archive's metadata, or via web-scraping the archive to identify pertinent variables. Due to the multilingual nature of the datasets, variables were translated using the Google Translate Application Programming Interface (API) and we cross-checked any ambiguous or unknown colloquial crop name against several sources (156, 157) and/or with colleagues who work in each region of interest. For each country in each data archive, we searched for variables that directly linked 'farm size' or 'plot area' with 'production' or gross 'plotted/cropped/planted/harvested area' by 'crop type'. If there were multiple eligible datasets available per country, we included the most recent year. Nearly all the source data were freely obtained and all are used according to their user agreements.

Table S1: Data repositories

Name	Region	Link
Accelerated Data Program	Global	http://adp.ihsn.org/country-activities
Africa Bank Group	Africa	https://www.afdb.org/en/knowledge/statistics
African Growth and Development Policy	Africa	http://www.agrodep.org/datasets
Consultative Group to Assist the Poor	Global	http://www.cgap.org/data
DataFirst	Africa	https://www.datafirst.uct.ac.za
Earthstat	Global	http://www.earthstat.org
Harvard's Dataverse	Global	https://dataverse.harvard.edu
Harvest Choice	Global	https://harvestchoice.org
International Food Policy and Research Institute	Global	http://library.ifpri.info/data
International Household Survey Network	Global	http://catalog.ihsn.org/index.php/catalog
Living Standards Measurement Study	Global	http://www.worldbank.org/en/about/unit/unit-dec
Prism	Oceania	http://pdl.spc.int/index.php/catalog
UNICEF Multiple Indicator Cluster Surveys	Global	http://mics.unicef.org/surveys
World Bank's microdata repo	Global	http://microdata.worldbank.org
World Food Program	Global	http://nada.vam.wfp.org/index.php/catalog
World Food Programme's Survey Data Portal	Global	http://nada.vam.wfp.org

Of the censuses that we included and had detailed sampling information (25 countries), 15 countries relied on either an exhaustive sampling design or a design that was exhaustive for farms with a set number of employees and/or annual revenue and a sample survey for smaller farms. Of the exhaustive censuses, there was a median response rate of 80%; the remaining censuses relied on stratified randomized sampling and applied resampling weights and expansion factors before making their aggregated data available (see dataset's metadata).

Farm size harmonization

We adjusted the tabulated census data to match the census data to the farm size classes that were reported in the WCA in order to enable consistent analyses across all countries. In some instances, census data farm size classes could simply be aggregated to match those reported in the WCA. In other instances, census data classes needed to be disaggregated into two or more WCA classes. For countries that had both tabulated census data and microdata available, the available area data in the microdata was aggregated into WCA classes, and the proportion represented by each class was used to distribute census data. For countries that had agricultural area by farm size class reported at the national level in the WCA, the proportion of area in each class was used to disaggregate subnational census data classes where necessary. For example, Paraguay reported a farm size class of 1-5 ha, whereas the WCA reported classes 1-2 ha and 2-5 ha. The total area in the 1-5 class was split between the two smaller classes based on their relative size, so 25% of area was assigned to the 1-2 ha class, and 75% of area was assigned to the 2-5 ha class. For all other countries, the simplest solution was to disaggregate classes that did not match based on the size class. There were instances where two different methods were used for the same country, for example in countries where the WCA only reported data in classes up to 100 ha. Figure S2 shows all reported farm size classes for tabulated census data (all European countries reported in Eurostat had the same classes, represented by the Europe category in Figure S2). The WCA classes, which were used in our analyses, are also shown. Corrections were made for the following countries:

Austria, Belgium, Brazil, Bulgaria, Croatia, Cyprus, Czech Republic, Denmark, Estonia, Ethiopia, Finland, France, Germany, Greece, Hungary, Iceland, India, Ireland, Italy, Latvia, Lithuania, Luxembourg, Malta, Montenegro, Netherlands, Norway, Paraguay, Poland, Portugal, Romania, Slovakia, Slovenia, South Africa, Spain, Sweden, Switzerland, United Kingdom, United States of America.

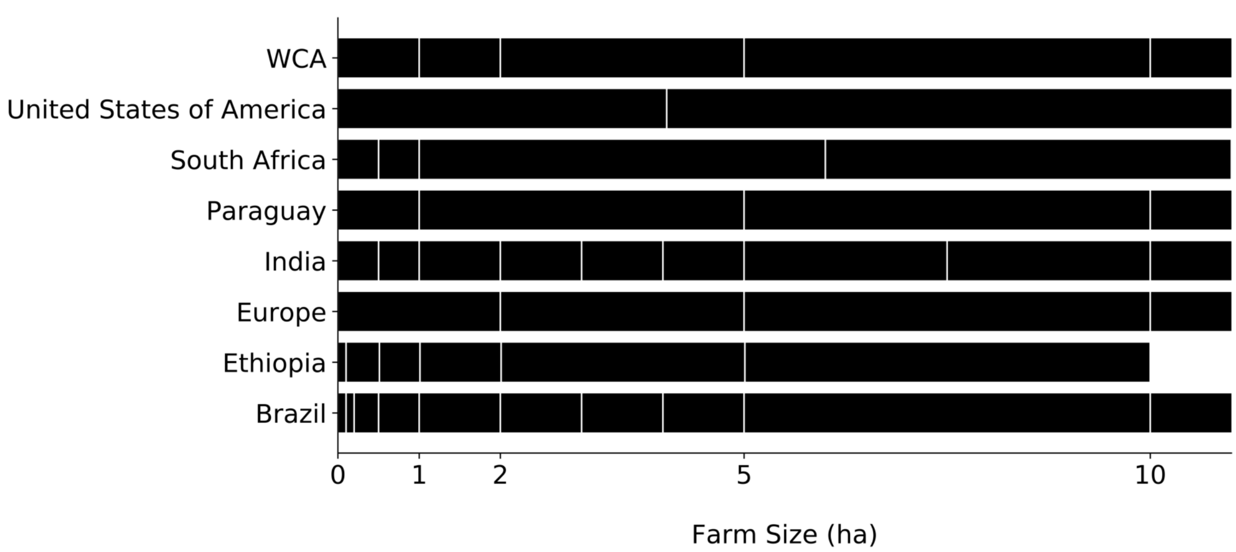

Figure S2: Farm size harmonization. Countries shown are where the given farm size classes were harmonized against the World Census of Agriculture (WCA) farm size classes. European countries from the Eurostat database had common farm size classes and are grouped together. Any country not shown contained directly matched farm size classes to the WCA. Since the majority of re-grouping occurred < 10 ha, the remaining farm size classes are not shown.

Construction of conversion factors

Conversion factors for kilocalories, fats, and proteins (in grams per capita) and for the percentage of each crop grown for food, animal feed, processed commodity, seed, and wastage due to transportation and storage (but not home consumption) were calculated using FAOSTAT.

FAOSTAT provides actual values for each of these variables at the national level per year with detailed definitions. For example, if a country produced soybeans in a given year, we took the ratio of the amount of soybean production allocated towards food divided by the total soybean production in that country to obtain the conversion factor for that country and year. We would repeat for feed, processed goods, seed, and waste, then apply these conversion factors to the amount of production each farm size produced per administrative unit in that country, and for each crop type. Hence, each estimate for these macro-nutrient and production variables assumes the national allocations are homogeneous across all administrative units and across all farm sizes. This is a largely untested assumption, and to our knowledge there are no sub-national datasets nor farm size specific datasets covering these variables, and therefore the bias introduced by it is unknown (unlike for other assumptions for which we were able to estimate bias, see Section 4). To enable future researchers to accommodate adjusting these conversion factors, we provide the actual amount of production per farm size per administrative unit in addition to the conversion factors and converted values.

Dataset descriptive statistics

Our dataset includes primary datasets ranging from 2001 to 2015, with a median year of 2013. It includes 55 countries, 45 of which have subnational resolution, 18 of which have fine scale (i.e., farm level) resolution.

Figure S3 shows the data's spatial resolution and distribution of the 154 unique crop species represented; on average (Mean), there were 30.8 crop species per country (Standard Deviation (SD) = 20.3). Crop species were paired to major commodity groups according to FAOSTAT definitions of cereals, fruit, oil crops, pulses, roots and tubers, tree nuts, vegetables, and other. Relying on the FAOSTAT classification has its limitations. For example, soy was classified as an oil crop, but it is also a pulse; therefore, this classification should be used as a guideline (Figure S3). Due to the aggregated nature of a large number of the sources used, we were only able to present gross agricultural area, not net agricultural area or the number of farmers by farm size class.

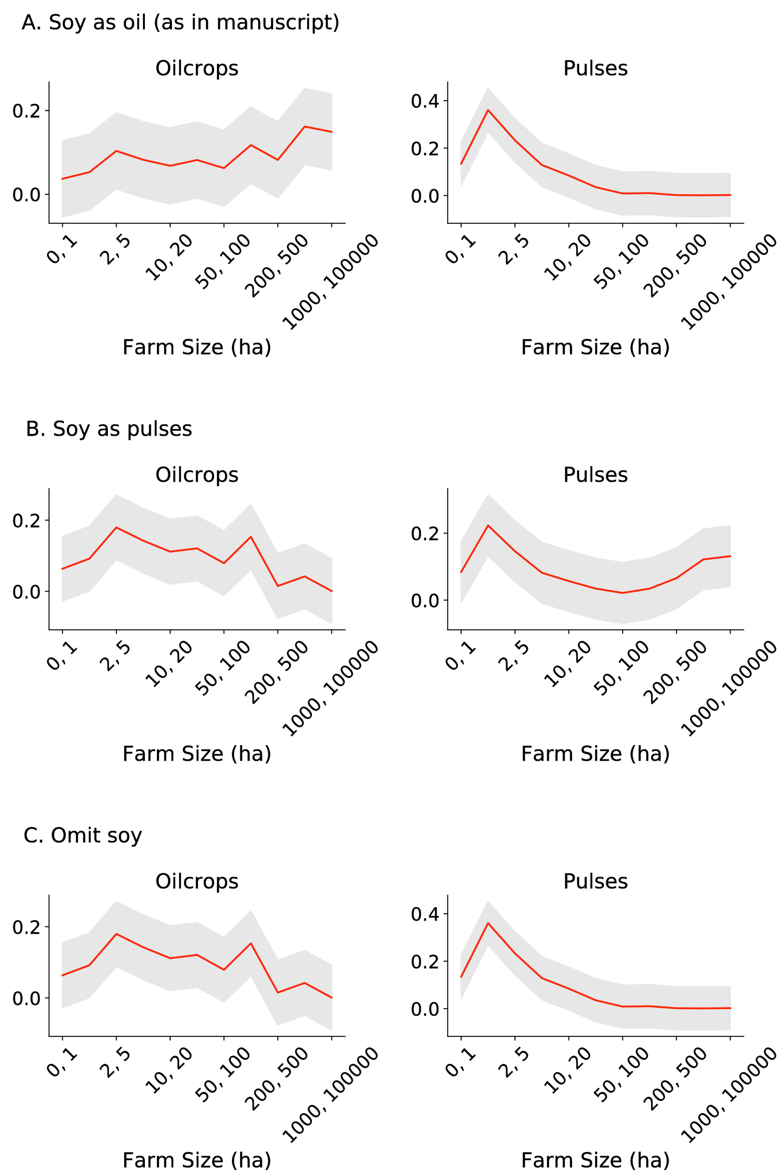

Figure S3: The effect of different classifications of soy on farm size distributions for oil crops and pulses. Soy was classified as an oil crop (Panel A as in our dataset and FAOSTAT), as a pulse (Panel B), or omitted (Panel C). The x-axis shows each farm size class (ha). The y-axis shows the percent of production. The red line is the average percent of production by farm size class. The gray line indicated 95% confidence intervals.

A.2.2. Key assumptions

Constant yields

For 33 countries in our dataset, representing 59.7% of the total production (in kcal), we could not find crop production by farm size, but we did find either gross cropped area, harvested area, planted area, or plot area by farm size per crop (Figure S4). For these data, we used FAOSTAT's national yield estimates for the given country, year, and crop to estimate production per farm size. This assumes that all farm sizes within a country had the same yields for a given crop and year. However, as there is a widely observed inverse yield to farm size relationship where smaller farms typically have higher yields (79, 83, 155), we explored how using a constant yield across farm sizes may bias our production estimates.

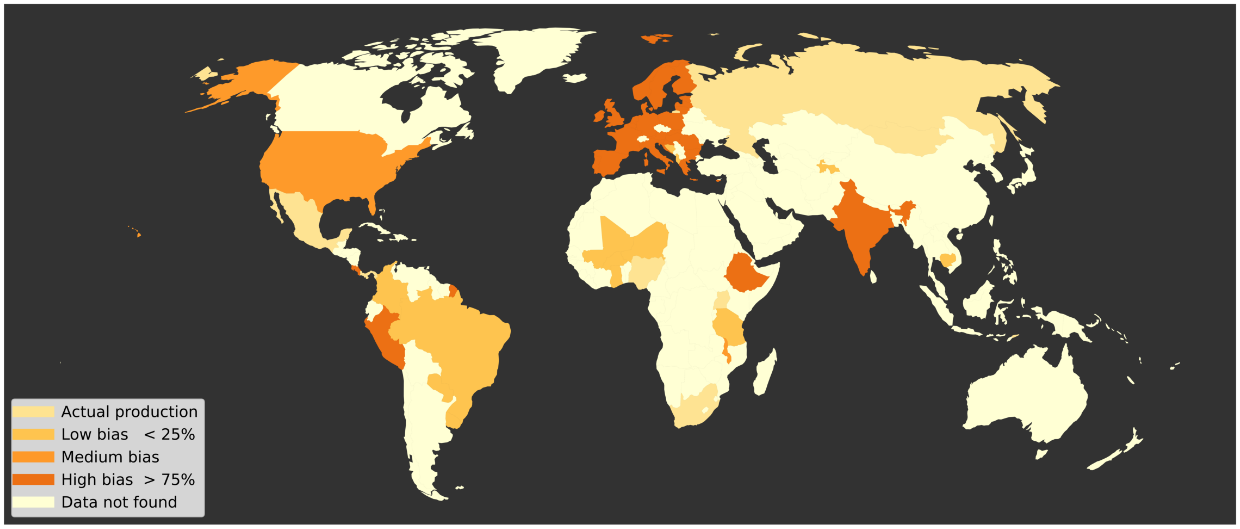

Figure S4: Map showing countries requiring assumption of constant yield across farm sizes. For many countries, our dataset contained a mix of actual production values and only area measurements per crops per farm size; percentages are given for each country according to how much of total crop production was calculated using constant yield assumption (indicated as percent bias in the legend). Darker orange indicates a greater percentage of the country's data was based on constant yields.

We tested the presence of a constant yield bias in eight countries for which we had both an area measurement (i.e., harvested, cropped, planted, or plot area) per crop per farm size and crop production by farm size measurement. For these countries, we regressed known production values against production values calculated from constant yields with countries and crop type as random effects, and we report the intercept and slope for this relationship to indicate the level of bias introduced by the constant yield assumption.

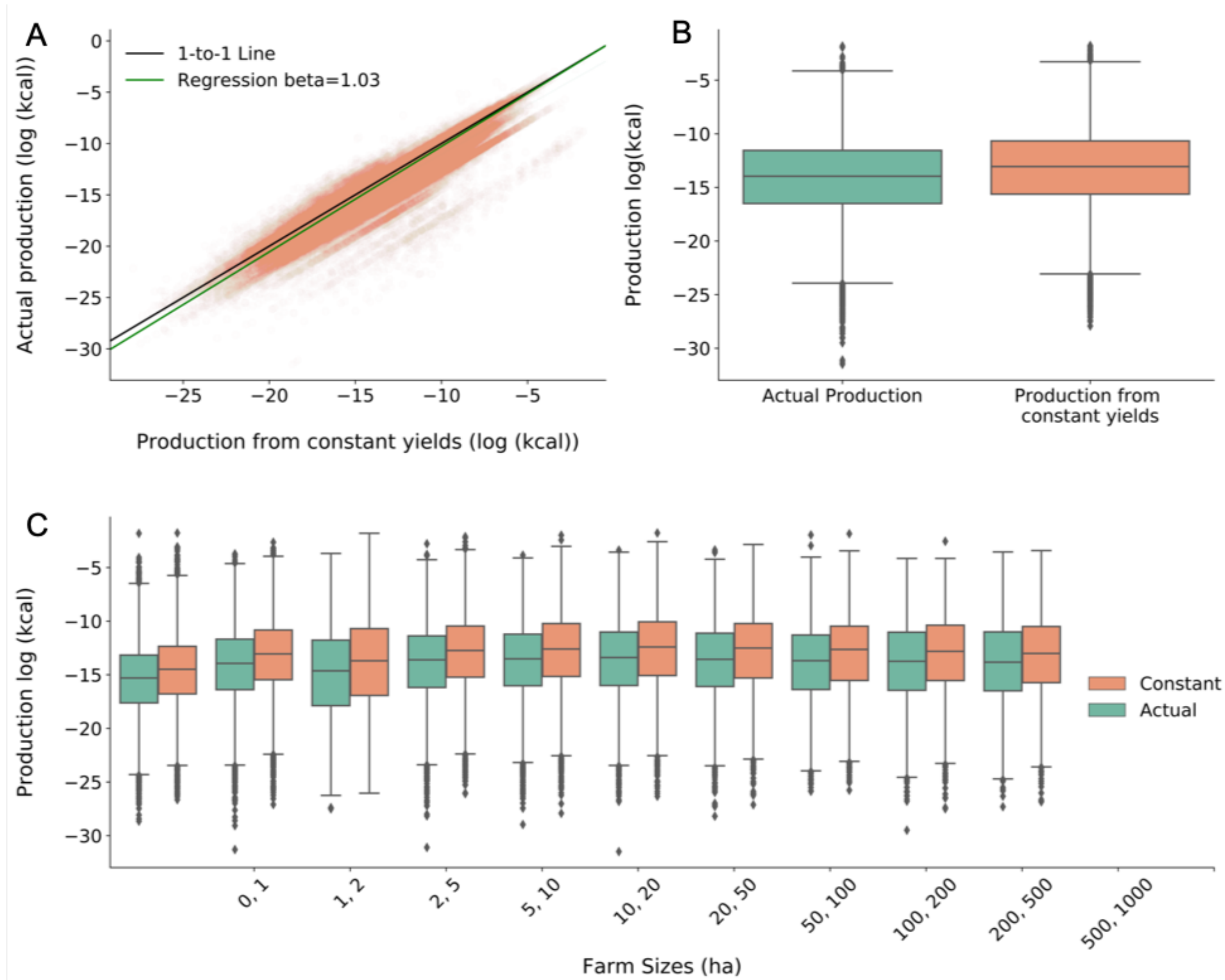

Figure S5: Verifying our constant-yield assumption through comparing production calculated using constant yields versus actual production for countries where we had both area and production data by farm size. A) Log-log plot between constant yield calculated production and actual production. Black line represents 1-to-1 line. Green line is the linear regression line when using constant yield derived production to predict actual production. B) Compares production using constant yields (orange) to actual (green) production on a log-scale, while C) shows this relationship for each farm size class.

Figure S5A is a log-log plot that shows a high correlation between production computed using constant yields and actual production. We used the natural log of production values to plot this due to long-tailed distributions in the data. We found that using constant yields slightly overestimates actual production for administrative units with smaller production but converges at administrative units with larger production (Intercept: -0.79, SE=0.11; Slope: 1.03, SE=0.001). This bias can be corrected for by predicting out of the model shown in Table S2. In Figure S5B, we also show boxplots to illustrate this overestimation for all farm size classes, and in Figure S5C we show the differences for each farm size. The plots indicate that overestimation of production from using constant yield is generally consistent across farm size classes.

Table S2: Constant yields at the national level were used to calculate production from cropping area at the sub-national level, then predict actual production. A mixed model was used to account for within country random effects.

Dependent Variable: Actual Production

	Coef.	**Std.Err.**	**95% CI**
Intercept	-0.786*	0.112	-1.005 to -0.567
Production from Constant Yields	1.028*	0.001	1.026 to 1.03
Group RE	4.771	0.484	
N Observations	95850		
N Groups	395		
BIC Full Model	212369.2		
BIC Without Constant Yields	455736.8		

Note: * = $p < 0.01$

Where country level yields were not available for certain crops and/or years, regional or global yields were used. Regional and global yields were used for 0.02% of all administrative units in our dataset (and had a Spearman rank correlation of 0.86 with the FAO country level yields) and so we expect them to have small effects on production values estimated across the sample. These are included in the constant yields assumption and the above bias analysis, but are denoted in the dataset for future researchers.

Calibrating with FAOSTAT

To calibrate our dataset with FAOSTAT we regressed our estimates of country production against theirs for matching crops and years. Our data consistency underestimates production relative to FAOSTAT (Intercept: 13.29, SE=1.52, and Slope: 0.99, SE=0.07; Figure S6). This relationship can be used to calibrate our data against FAOSTAT for future researchers interested in using this data. As we used the exact matching of crop lists with the FAO, this is perhaps surprising. It is possible that some of this variation represents differences in survey instruments since we have included different datasets from what FAOSTAT included since we needed to have access to crop production by farm size and FAOSTAT did not require this cross-tabulation. Another way of looking at this discrepancy is that our dataset provides an independent, and transparent, estimate of the amount of crops produced by different countries across the world.

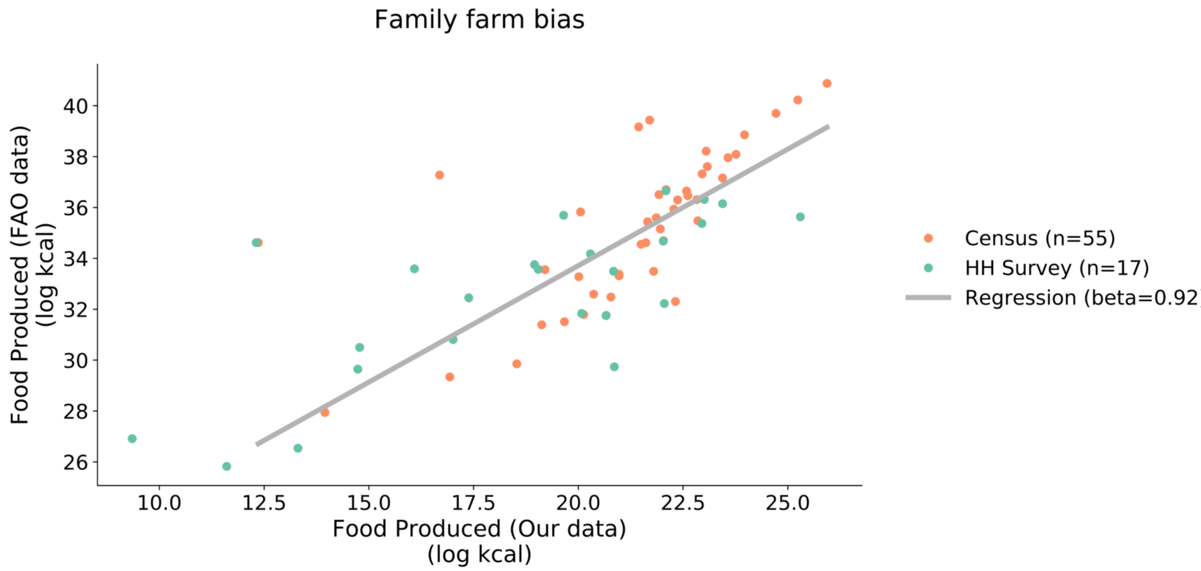

Figure S6: Log-log plot comparing FAOSTAT production values (summed kcal crop equivalents per country) to our dataset with and without household surveys. Household surveys are in green, census data are in orange. The simple linear regression line shows the relationship between the summed production values for countries in our dataset with their FAOSTAT summed production values.

Family farms bias

For 17 countries in our dataset, representing 22.5% of the total production (in kcal), we could not find agricultural census data, but we did find nationally representative (often with sub-national resolution) agricultural household surveys (Figure S1). One bias that stems from household surveys is that they only capture family farms, which are often associated with smaller farms. The household surveys miss non-family commercial enterprises and thus do not represent the full population of farms in a country. A proper test of the bias introduced by use of household surveys would require both census and household survey data for the same countries, which we did not have access to for the countries in our dataset and they covered different ranges and magnitudes of production (e.g. with household survey data covering countries with smaller aggregate production; see Figure S6).

Plot size as a farm size proxy

For 8 countries in our dataset, representing 4.8% of the total production (in kcal), farm size was not explicitly reported, so we calculated a proxy farm size using the sum of either harvested, cropped, planted, or plot area (Figure S7). This assumption may influence estimates of global crop production by farm size by underestimating farm areas in some farm size classes, because the aggregation process did not capture all fallow plots, water sources, unused areas, and on-farm structures. We think the main effect of this would be to introduce noise into the production by farm size signal (by mixing data using the field size proxy with real farm sizes). Due to data constraints, we were not able to explore how much noise this introduced. It does stand to reason that larger fields need to belong to larger farms, but it is unclear whether smaller fields are part of a large farm with several small fields or part of a small farm. However, because these countries represent less than 5% of the total production covered in our dataset, they do not greatly influence gross estimates of crop production by farm size estimated from these data. When the 8 countries we used a proxy indicator for farm size are omitted from the dataset there was minimal influence on the distribution of food production by farm size (mean absolute difference=0.26; SD=0.19).

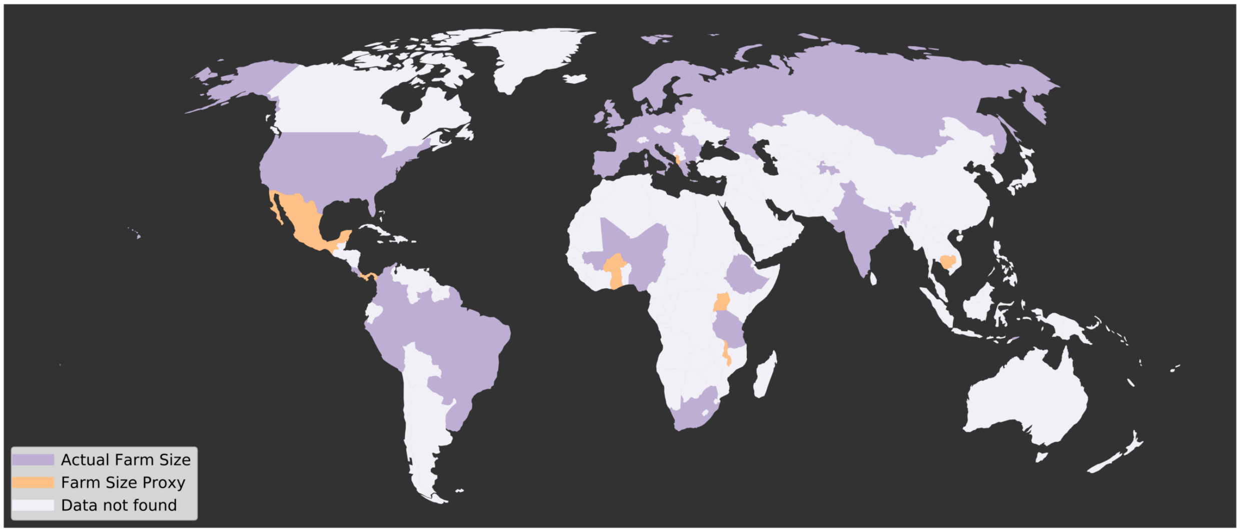

Figure S7: Map showing direct farm size data (purple) or farm size proxy (orange) at the country level.

Regional bias

Our dataset accounts for around 51% of the total global harvest area, with representation across country types (e.g., spatial and economic). However, since our dataset is a convenience sample, we were not able to control for spatial coverage nor the countries included, and there were large data gaps for Australasia and Asia Figure S8).

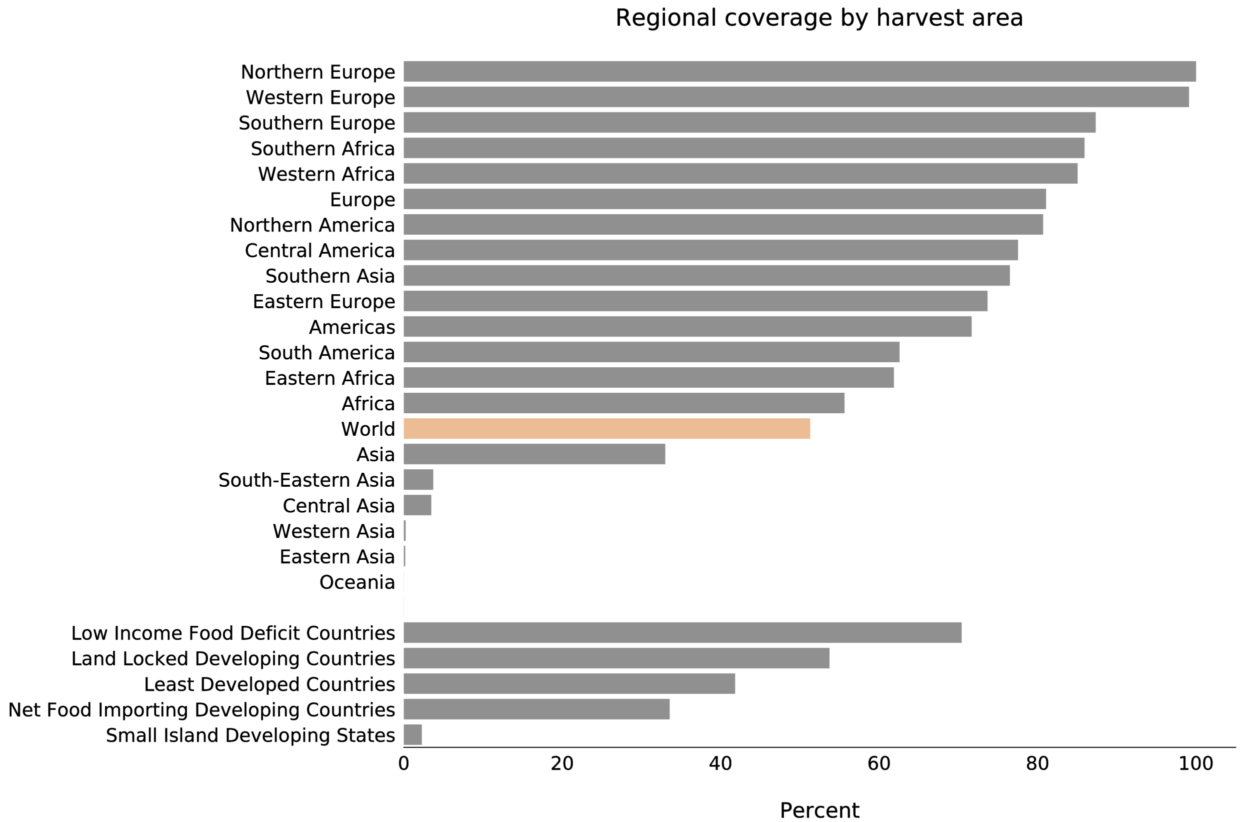

Figure S8: Dataset's percent of harvest area by region or economic status compared to global coverage in orange. Harvest area per region calculated from FAOSTAT.

An important question for researchers interested in this dataset is how much the global estimates of crop production by farm size are influenced by the omission of particular countries. While this coverage error is difficult to compute directly, we can explore how sensitive global estimates are to any one country included in the dataset. To do this we re-computed a leave-one-out jackknife statistic, shown in Figure S9. The vertical black line is the mean kilocalories (kcal) of food produced for a given farm size class when no countries were omitted. Each blue dot represents the mean when a corresponding country was omitted. If a country is to the left of the black line it lowers the global average. The vertical lines are the upper and lower quartiles for food production. For each plot, we labelled four countries as examples, but all countries are present.

There is substantial variation when a country is omitted indicating that countries' farm size distributions can heavily influence the global averages (see Tables 3-5 for per country distributions of gross agricultural, total production (kcal), and food production (kcal)). This high variation in the percentage of food produced in different farm size classes indicates that the relationship between farm size and food production is highly contextual; Figure S10 shows two examples, South Africa and Germany.

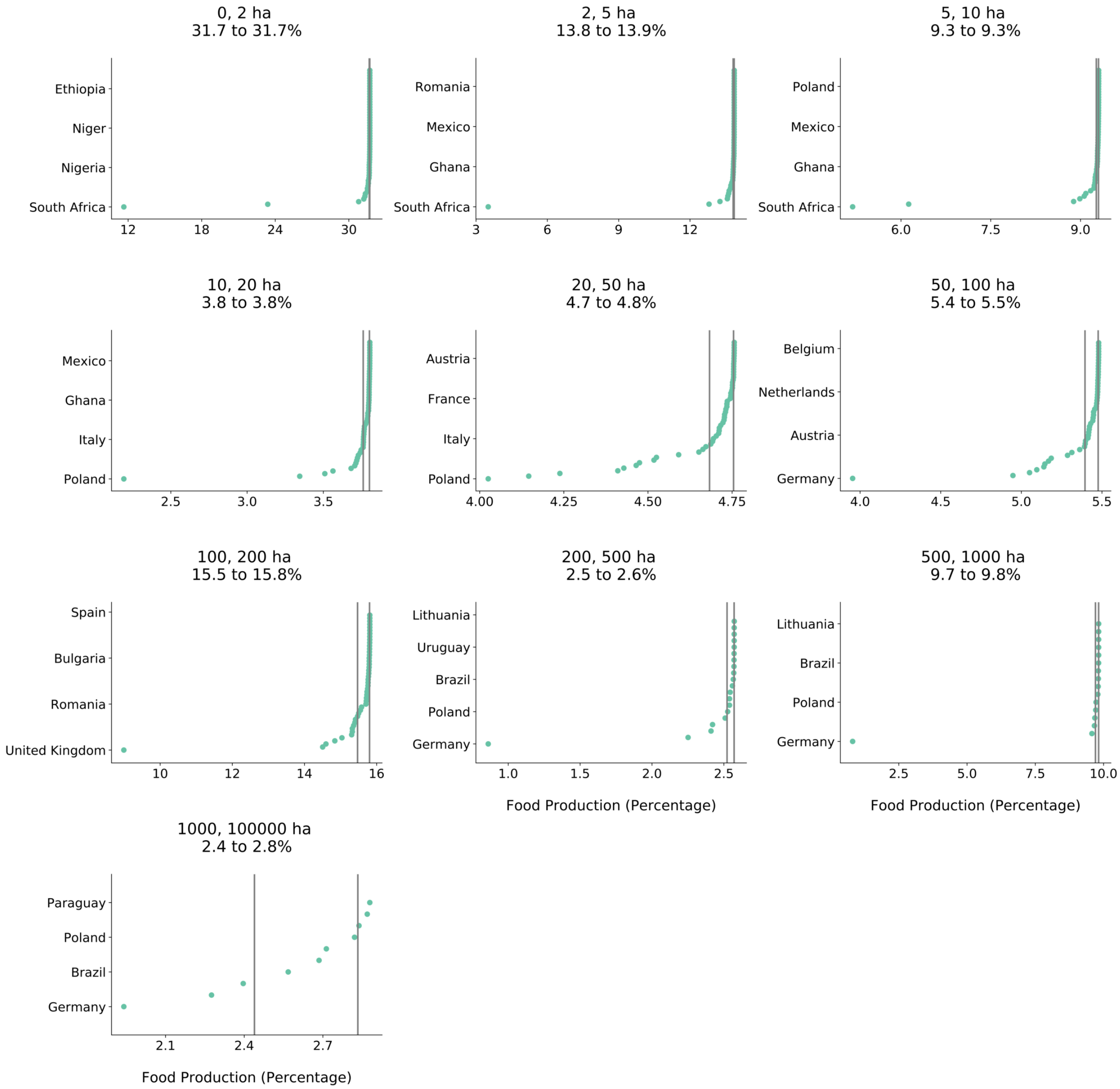

Figure S9: Jackknife plots per farm size to estimate country level bias. Grey lines indicate upper and lower quartiles of global production, and green points refer to the global mean if the country was omitted.

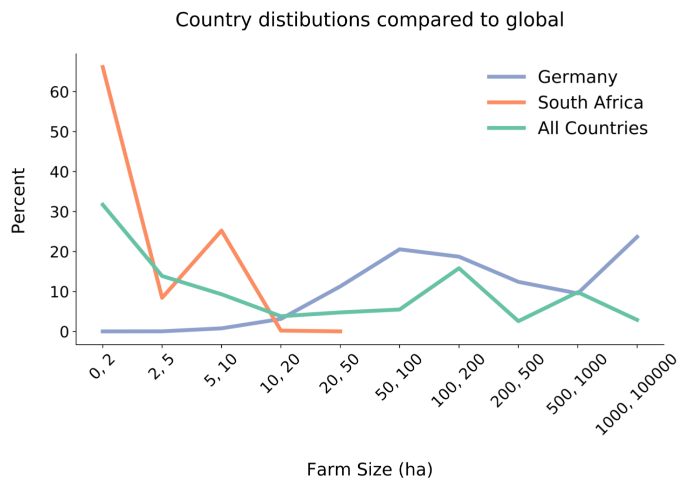

Figure S10: Two examples of countries that deviated from the global distribution of total crop production by farm size: Germany (purple) and South Africa (orange) have different distributions than the global average (green).

A.3. Global production statistics

Table S3: Crop production by allocation type (kcal 10e9) and gross agricultural area (ha2 10e9) per farm size class.

Farm Size (ha)	Feed	Food	Other	Processing	Seed	Waste	Area
0, 1	9.5	65.0	28.6	2.6	2.2	4.0	15.5
1, 2	24.7	76.6	22.2	4.3	2.5	5.8	15.7
2, 5	11.5	61.9	27.4	4.7	2.5	4.7	9.5
5, 10	15.1	41.5	12.6	3.9	1.7	3.5	11.4
10, 20	5.4	17.0	5.8	3.7	1.2	1.6	10.9
20, 50	8.1	21.2	4.2	7.9	1.7	1.5	10.5
50, 100	10.1	24.5	3.7	10.9	2.1	1.8	10.0
100, 200	23.4	70.6	8.7	27.1	6.8	6.1	9.6
200, 500	8.4	11.5	3.5	4.6	1.0	0.9	10.1
500, 1000	20.1	43.9	7.1	15.0	3.8	2.0	10.0
1000, 100000	6.6	12.9	3.6	7.4	1.8	2.6	15.9

Table S4: Species richness distributions by farm size class.

Farm Size (ha)	Mean Richness (%)	95% CI Low	95% CI High
0 to 1	19.92	15.38	24.46
1 to 2	22.76	18.22	27.3
2 to 5	12.68	8.14	17.22
5 to 10	16.66	12.12	21.2
10 to 20	14.01	9.47	18.55
20 to 50	11.92	7.38	16.46
50 to 100	8.73	4.19	13.27
100 to 200	7.5	2.96	12.05
200 to 500	6.73	2.19	11.27
500 to 1000	6.36	1.82	10.9
1000 to 100000	7.05	2.51	11.59

Table S5: Production (kcal 10e10) by crop type per farm size class.

Farm Size (ha)	Cereals	Fruit Incl Melons	Oilcrops Primary	Other	Pulses	Roots and Tubers	Treenuts	Vegetables Primary
0 to 1	7.10	0.07	0.02	0.10	0.04	0.57	0.00	0.11
1 to 2	10.04	0.07	0.02	0.09	0.11	0.58	0.00	0.09
2 to 5	6.85	0.12	0.05	0.14	0.07	0.68	0.00	0.10
5 to 10	5.40	0.07	0.04	0.15	0.04	0.45	0.00	0.06
10 to 20	1.76	0.05	0.03	0.19	0.03	0.34	0.00	0.07
20 to 50	2.08	0.03	0.04	0.63	0.01	0.24	0.00	0.15
50 to 100	2.27	0.01	0.03	0.99	0.00	0.24	0.00	0.22
100 to 200	7.04	0.02	0.05	2.77	0.00	0.39	0.00	0.28
200 to 500	1.84	0.00	0.04	0.44	0.00	0.08	0.00	0.04
500 to 1000	5.00	0.00	0.07	1.80	0.00	0.25	0.00	0.04
1000 to 100000	1.99	0.01	0.07	0.73	0.00	0.10	0.00	0.08

Table S6: Macronutrient production by farm size class. In grams per capita (10e9) of carbohydrates, fats, or protein per farm size class.

Farm Size (ha)	Carbohydrates	Fats	Proteins
0, 1	112.34	0.78	6.31
1, 2	78.84	0.56	4.42
2, 5	90.49	0.66	5.10
5, 10	34.18	0.26	1.90
10, 20	14.25	0.11	0.77
20, 50	7.89	0.06	0.44
50, 100	2.36	0.02	0.12
100, 200	1.44	0.01	0.07
200, 500	1.71	0.01	0.10
500, 1000	1.56	0.01	0.10
1000, 100000	7.40	0.06	0.46

Appendix B Chapter 3 supplemental information

B.1. Detailed description of data and methods

B.1.1. Literature survey

Studies were coded at the observational level to analyze multiple crops, years, and locations per study; studies had multiple observations if they separately reported different crops, years, and/or locations per outcome variable. For each observation, we recorded mean farm size, crop(s), non-crop specie(s) for biodiversity, the scale of analysis (farm or landscape), and location. The main conclusions were categorically coded as "vote-counts", where an increase in farm size was associated with a decrease, increase, or null relationship to the variable of interest (we found no non-linear results in the literature that met our inclusion criteria). For yield, resource-use efficiency, and profit we extracted several additional variables to calculate pooled effect sizes: regression model coefficients and standard errors; the type of model used; mean and units of outcome and response variables; sample size per observation; type of metric used (e.g., yield defined as either kg/ha or value output/ha; resource-use efficiency measured as technical efficiency (using either Cobbs-Douglas or stochastic frontiers); efficiency factors included in resource-use efficiency studies (e.g., seed, irrigation, fuel, labor, etc.); control variables in the regression equation (e.g., for yield studies, the type of production system, labor, land heterogeneity, and/or institutional factors were common controls; while credit access, extension access, and membership to farmer cooperatives were common controls in resource-use efficiency studies). Due to finding a limited number of studies that directly measured farm size and GHG emissions per unit output, we leveraged the Clark and Tilman meta-analysis database containing 742 agricultural life-cycle analysis (LCA) observations from 152 unique studies (108); we coded observations that reported average farm size to construct a dataset containing crop species, GHG emissions per unit output (in CO2 equivalents), average farm size, and sample size for 100 observations (11 studies) that met our inclusion criteria. Similarly, there were limited crop diversity studies that met our criteria. To supplement these findings, we report the results of our recent study that measured crop diversity, in terms of relative species richness, and farm size across 55 countries and 154 crops (29).

B.1.2. Synthesis of results

Equation 1 shows the CLMM's general form, where $P(Y_i \leq j)$ is the cumulative probability of the i-th observation falling into the j-th category (negative, null, or positive relationship to farm size) (120). θ_j is a threshold parameter of designating cut-points between the three ordinal categories; hence, the difference between $\theta_{negative} - 0$, $\theta_{positive}$ - $\theta_{negative}$, and $1 - \theta_{positive}$ represent the average probabilities of a negative, null, or positive relationship, respectively. $X\beta_i$ is a matrix of all fixed effects and their slopes, in which $\exp(\beta)$ represent an odds ratio of the event $Y \geq j$. U_i is a matrix of the random effect terms; all random effects are crossed. We bootstrapped each model to calculate the 95% confidence intervals of the average probabilities of a negative, null, or positive relationship.

Equation 1 $\quad \text{logit}(P(Y_i \leq j)) = \theta_j - X\beta_i - U_i$

Equation 2 shows the general form of the robust linear mixed-effects model used for the GHG emissions analysis where we set location and crop type as random effects. Where y is the standardized regression coefficient per observation (i) (136). $X_i\beta$ is a matrix of fixed effects; Z_ib_i are the random intercepts; ε_i is the error term. The model calculates an average fixed effects intercept by random intercepts. We bootstrapped the models to extract the fixed effect intercept, then calculated the median effect and 95% confidence intervals. To calculate the random effects meta-regressions for yield, resource efficiency, and profit, we relied on the R package "metafor" (122). These meta-regressions are near identical to equation 1, except they do not estimate the standard errors, rather we used fixed variances from the extracted standard errors from each study.

Equation 2 $y_i = X_i\beta + Z_ib_i + \varepsilon_i$

B.2. Supplemental figures

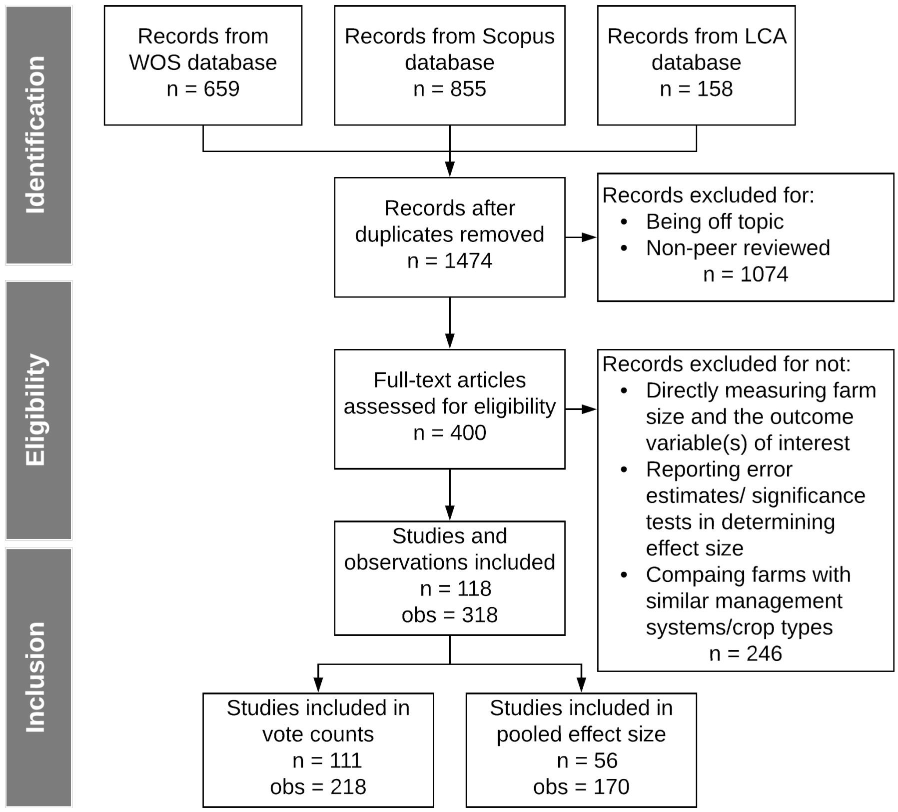

Figure S12: PRSIMA diagram of the data identification, eligibility, and inclusion process. Web of Science (WOS) and Scopus article databases were queried, the results were combined and duplicates were removed. Clark and Tillman's (2017) dataset of life-cycle analysis (LCA) studies were additionally coded if the study included farm size summary statistics. A few LCA studies pertained to other variables, beyond GHG emissions, and were included in the vote counts where appropriate.

Similarly, a few resource-use efficiency studies found via WOS and Scopus contained GHG information that further supplemented the LCA dataset.

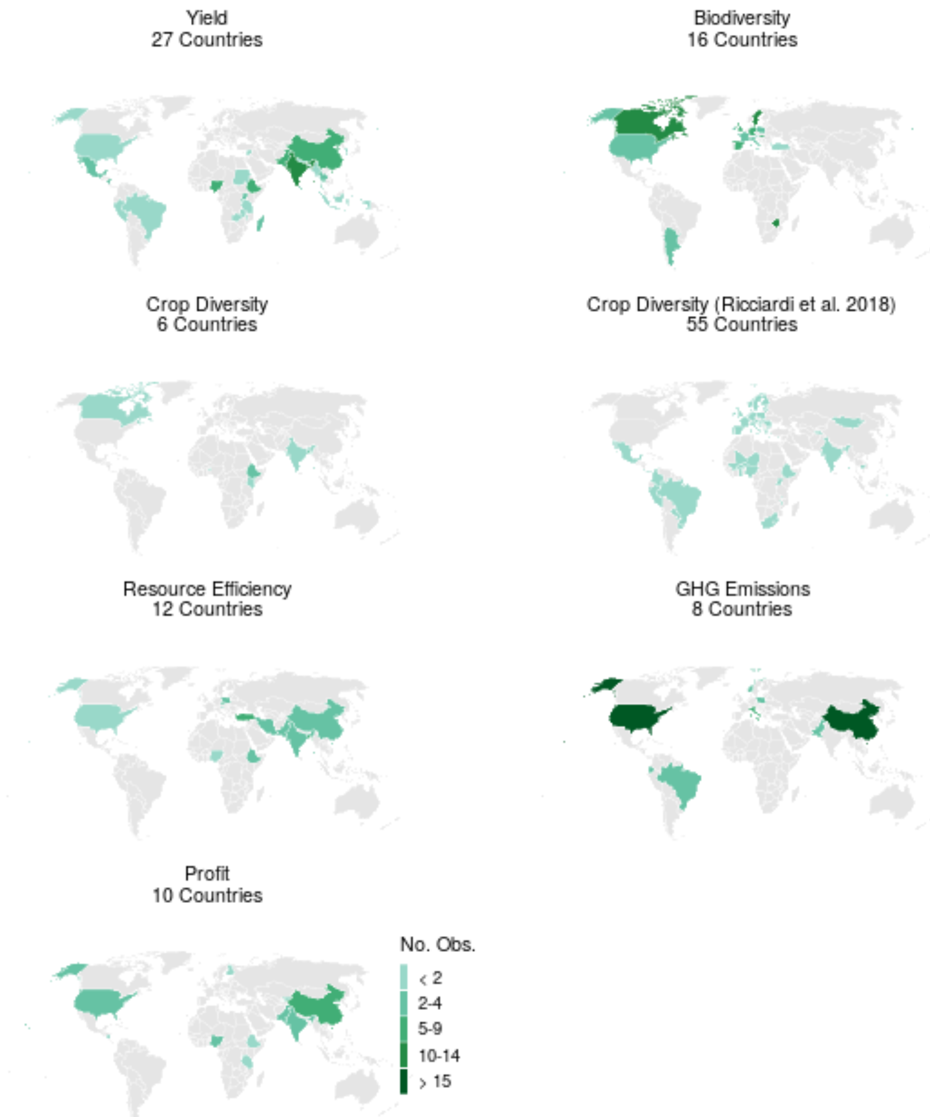

Figure S13: Number of observations per country per variable.

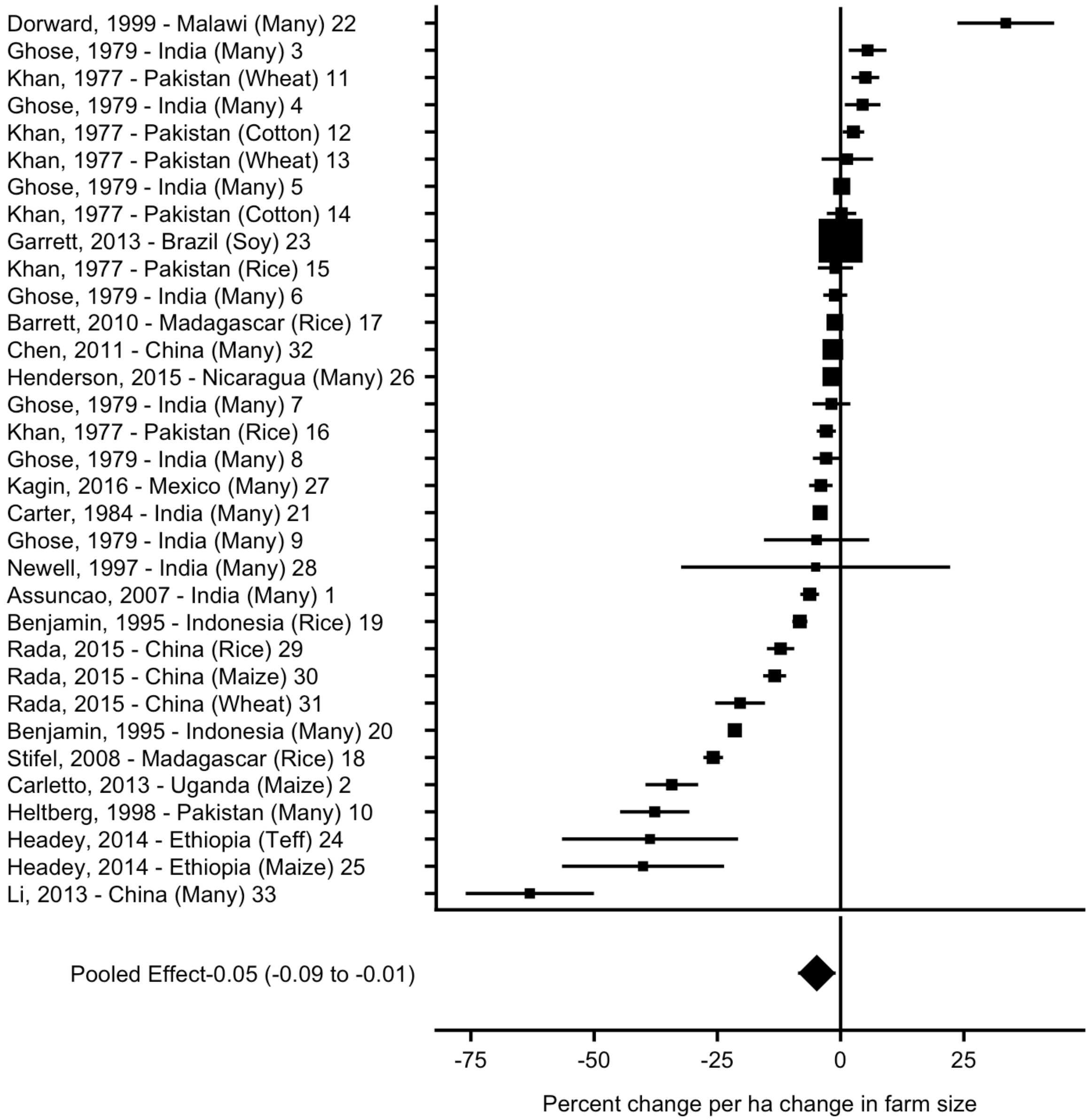

Figure S14: Forest plot for yields, where observations are in standardized form and 95% CI are given. The size of each point estimate relates to the inverse standard error. The pooled effect and 95% CI are given in the lower plot.

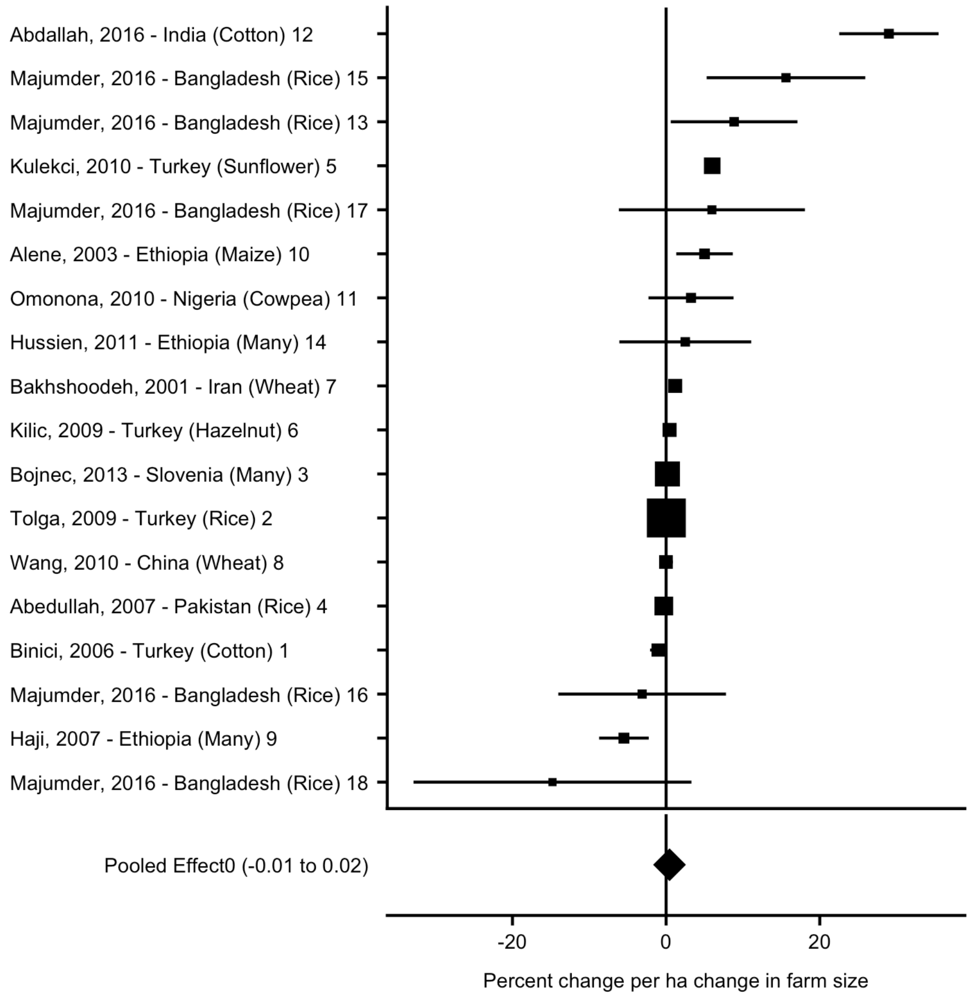

Figure S14: Forest plot for resource efficiency, where observations are in standardized form and 95% CI are given. The size of each point estimate relates to the inverse standard error. The pooled effect and 95% CI are given in the lower plot.

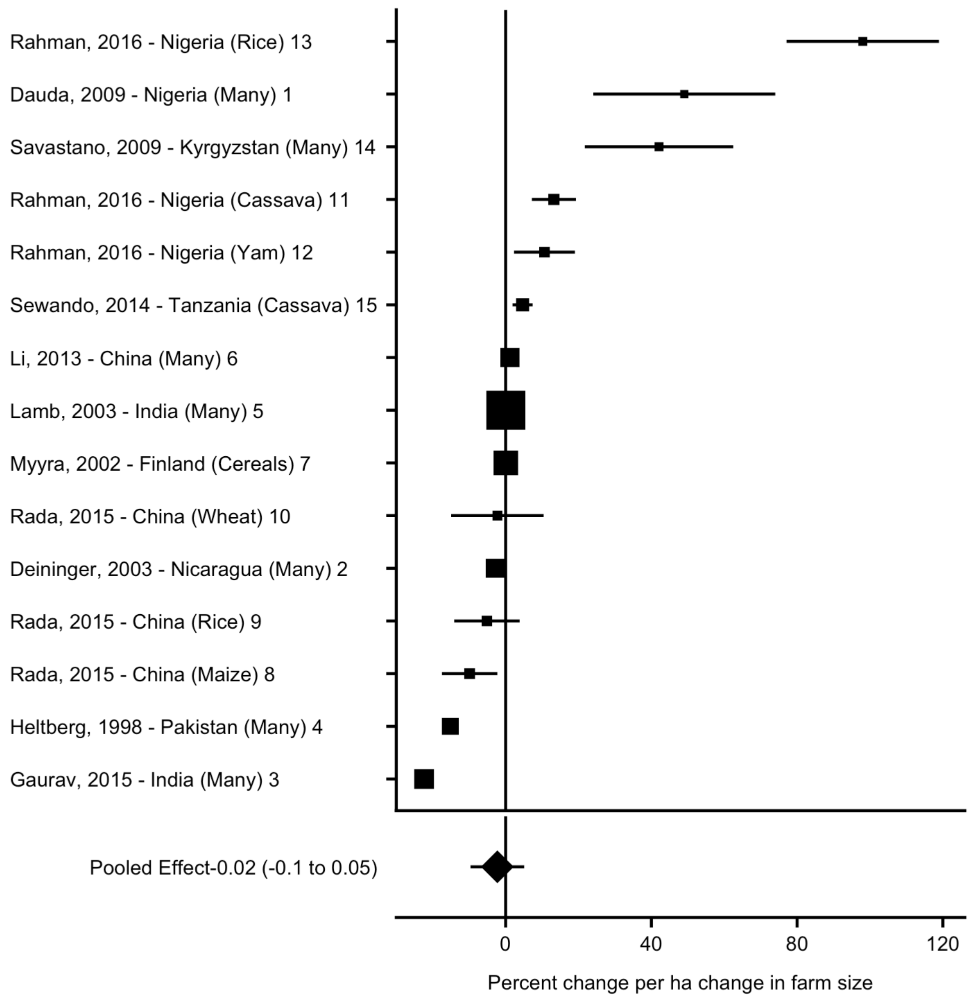

Figure S15: Forest plot for profitability, where observations are in standardized form and 95% CI are given. The size of each point estimate relates to the inverse standard error. The pooled effect and 95% CI are given in the lower plot.

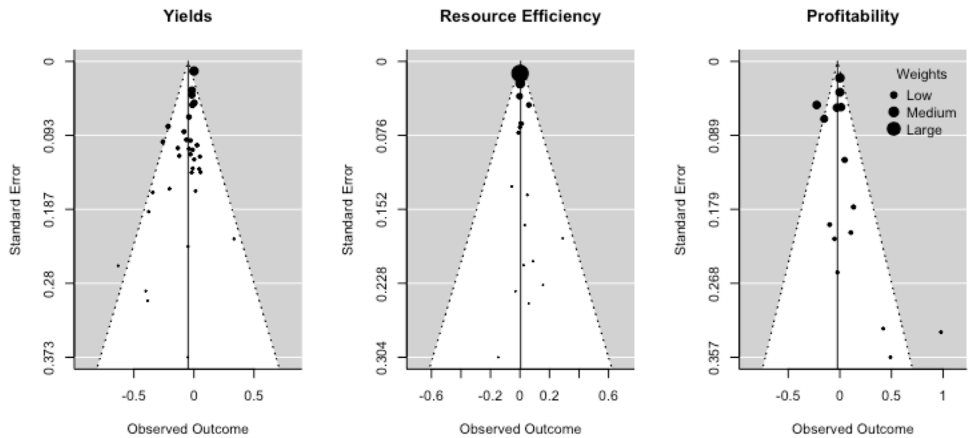

Figure S17: Funnel plots to identify bias for observations included in the meta-analyses, where the observed outcome is plotted against the standard error. Publication bias would result in non-symmetric plots.

B.3. Supplemental tables

Table S7: Boolean search terms per variable and number of articles returned.

Variable	Search Booleans	No. Articles Web of Science	Scopus
Yields	("Farm size*" OR "Field size*") AND (yield* OR invers*) AND (agricult* OR crop*)	278	423
	("Farm size*" OR "Field size*" OR "plot size*") AND (invers*) AND (agricult*)	56	79
	("Farm size*" OR "Field size*") AND (yield*) AND (invers* OR productivity) AND (agricult* OR crop*)	75	101
Biodiversity	("Farm size*" OR "Field size*") AND (biodivers* OR divers*) AND (environment* OR ecosystem* OR landscape* OR sharing)	156	198
	("Farm size*" OR "Field size*") AND (biodivers* or divers) AND (environment* OR ecosystem* OR landscape* OR sharing) AND (intensification)	41	34
	("Farm size*" OR "Field size*") AND (deforestation)	19	17
Crop Diversity	("Farm size*" OR "Field size*") AND ("crop divers*" OR "diet divers*" OR "in situ")	70	83
	("Farm size*" OR "Field size*") AND ("monocrop*" OR "intercrop*")	22	32
	("Farm size*" OR "smallhold*") AND ("crop divers*" OR "diet divers*" OR "in situ") AND ("size*")	42	48
	("Farm size*" OR "smallhold*") AND ("crop divers*" OR "diet* divers*" OR "in situ*")	136	174
Resource Efficiency	("Farm size*") AND ("input*") AND (use OR efficienc*)	107	263
	("Farm size*") AND ("energy efficienc*")	11	23
	("Farm size*" OR "Field size*") AND (input*) AND (use OR efficienc*) AND (agricul*)	126	187
	("Farm size*" OR "Field size*") AND ("energy efficienc*") AND (agricul*)	5	13
Profit	("Farm size*" OR "Field size*") AND ("profit*") AND ("agricult*" OR "crop*")	131	180
	("Farm size*" OR "Field size*") AND ("inverse*" OR "allocative efficienc*" OR "economic efficienc*" OR "viability" OR "viable") AND ("agricult*" OR "crop*")	105	159
	("Farm size*" OR "Field size*") AND ("inverse*" OR "allocative efficienc*" OR "economic efficienc*")	229	261
	("Farm size*" OR "Field size*") AND ("inverse" OR "allocative efficienc*" OR "economic efficienc*" OR "viability" OR "viable") AND ("agricult*" OR "crop*")	144	308
GHG Emissions	All data was from Clark and Tillman's (2017) LCA database available in their supplemental material.		

Table S8: Summary statistics.

Variable	Group	Finding	Median Year	No. Obs	No. Studies
Yield		Effect Size	1998	32	18
		Negative	1998	43	26
		Null	1984	16	9
		Positive	1979	10	6
Biodiversity	Field	Negative	2005	32	13
		Null	2012	14	8
		Positive	2012	6	4
	Landscape	Negative	2010	22	10
		Null	2012	10	6
		Positive	1992	3	2
Crop Diversity		Negative	2015	3	3
		Null	2015	1	1
		Positive	2011	4	4
Resource Efficiency		Effect Size	2010	19	15
		Negative	2007	6	6
		Null	2011	11	9
		Positive	2013	17	14
GHG Emissions		Effect Size	2016	100	11
Profitability		Effect Size	2014	15	11
		Negative	2014	6	6
		Null	2015	3	2
		Positive	2013	11	9

Table S9: CLMM Regression probabilities that a study finds a negative, null, or positive relationship between farm size and the variable of interest.

Variable	Relationship	No. Obs.	Mean	Lower 95%	Upper 95%
Yield	Negative	43	78.74	57.77	100.00
	Null	16	18.10	0.00	36.29
	Positive	10	3.17	0.00	8.06
Biodiversity	Negative	54	76.90	60.94	98.82
	Null	24	21.60	1.17	35.97
	Positive	9	1.50	0.00	4.27
Resource Efficiency	Negative	6	14.05	7.02	19.23
	Null	11	36.24	26.92	47.67
	Positive	17	49.71	39.29	58.39
Profit	Negative	6	21.78	0.00	39.71
	Null	3	22.24	0.00	55.41
	Positive	11	55.97	23.10	100.00

Table S10: Mixed effects regression output of effect size per variable with 95% confidence intervals.

Variable	Subset	No. Obs.	Mean Effect	Lower 95%	Upper 95%
Yields	All	33	-0.05	-0.09	-0.01
	Institutional not controlled	26	-0.05	-0.11	0.01
	Institutional controlled	7	-0.05	-0.16	0.06
	Labor not controlled	17	-0.06	-0.13	0.01
	Labor controlled	16	0.01	-0.08	0.10
	Management not controlled	16	-0.03	-0.09	0.02
	Management controlled	17	-0.04	-0.12	0.04
	Institutional controlled (full model)	7	-0.06	-0.19	0.07
	Labor controlled (full model)	16	0.02	-0.11	0.16
	Management controlled (full model)	17	-0.05	-0.17	0.08
Resource Efficiency	All	18	0.00	-0.01	0.02
	Cooperative not controlled	14	0.00	-0.03	0.03
	Cooperative controlled	4	0.00	-0.04	0.04
	Credit not controlled	8	0.00	-0.02	0.02
	Credit controlled	10	0.06	-0.02	0.13
	Extension not controlled	7	0.00	-0.02	0.02
	Extension controlled	11	0.05	-0.02	0.13
	Cooperative controlled (full model)	4	0.00	-0.04	0.04
	Credit controlled (full model)	10	0.02	-0.31	0.36
	Extension controlled (full model)	11	0.03	-0.30	0.36
GHG	All	100	-0.04	-0.10	0.02
Profitability	All	15	-0.02	-0.10	0.05

Appendix C Chapter 4 supplemental information

C.1. Supplemental figures

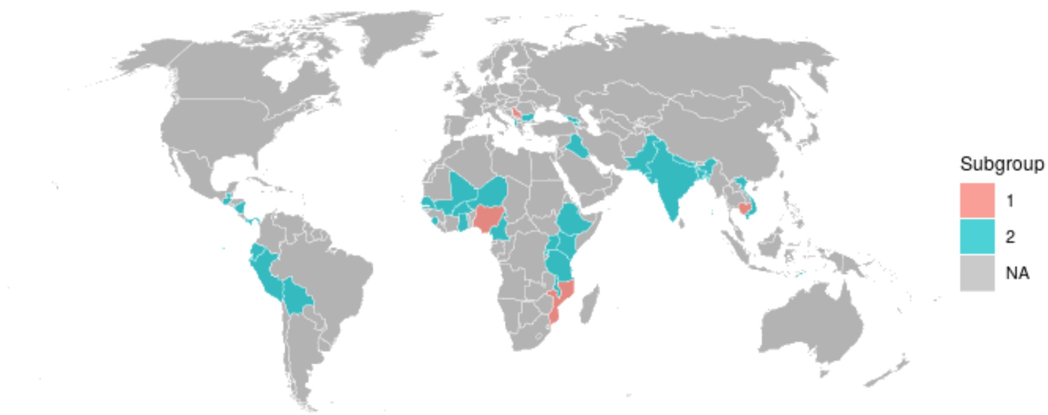

Figure S18: Map of the different subgroups we use for our analysis. Subgroup 1 (orange and green) contains 34 countries, subgroup 2 (only green) contains 30 countries. Each subgroup analysis uses a different selection of predictor variables. Subgroup 1's variables are: farm area (actual and relative), relative economic size, and the percent of female labor. Subgroup 2 includes all subgroup 1 predictors plus: percent family labor and percent subsistence. Countries in gray, indicate no available data.

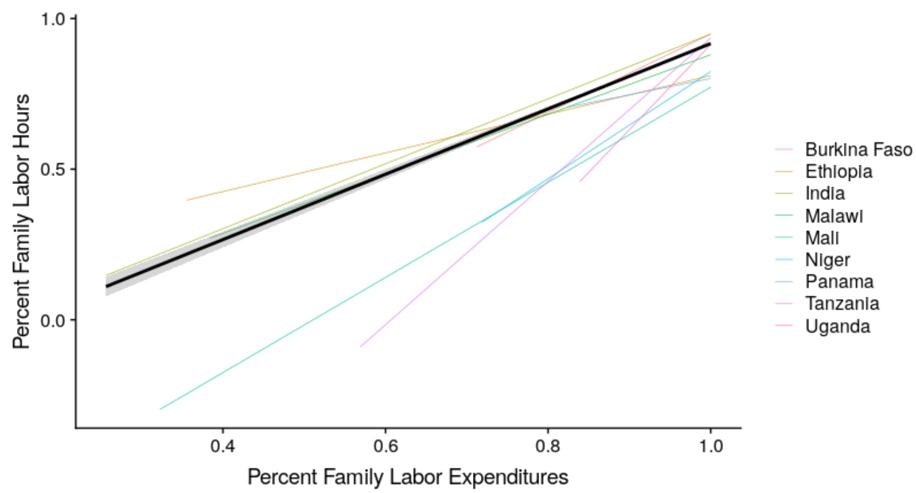

Figure S19: We defined out family labor variable by one minus the percent of labor expenditure over total on-farm expenditures, which assumed that family labor was unpaid. We used this definition because it allowed us to measure 30 countries. We tested this definition against the number of days of family labor worked per year over the total number of labor days worked per year per farm for nine countries where we had these variables (this would have been an ideal definition for percent family labor). We see a very strong linear relationship between our proxy variable and the actual percent family labor.

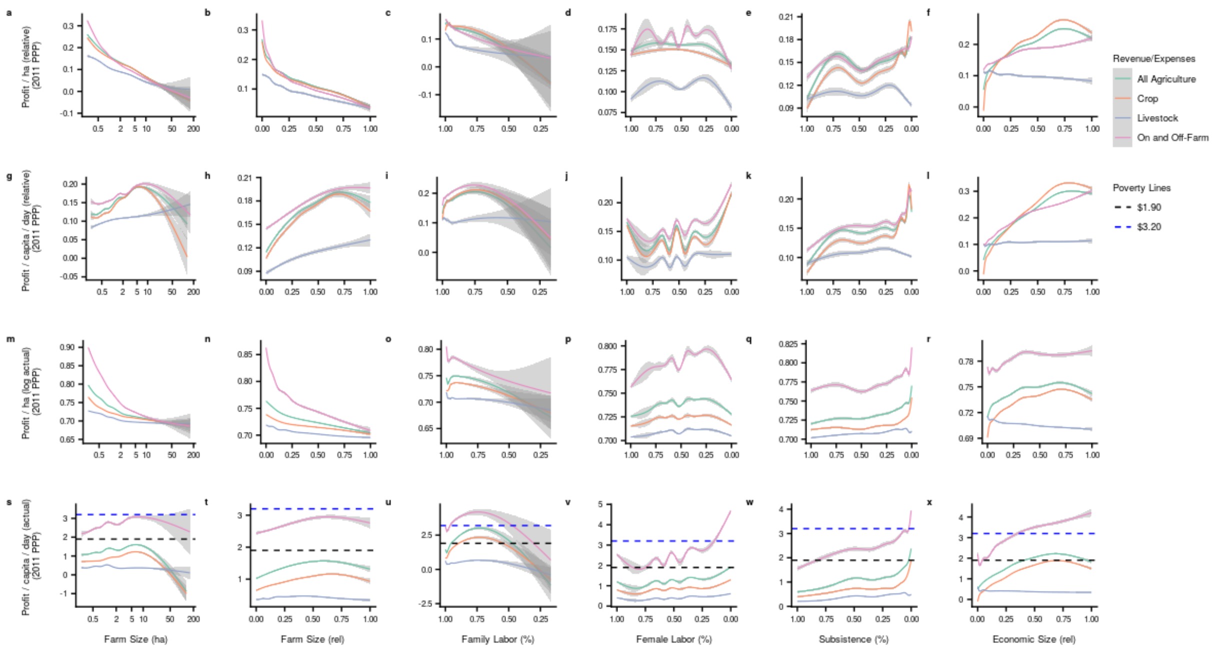

Figure S19: Bivariate relationships of each definition of smallholder with productivity (profit per ha) and income (profit per person in the farming household) in country relevant terms (upper two rows) and real terms (2011 USD PPP). Family labor, female labor, and subsistence level have reversed x-axis to enable easier visual comparison across smallholder definition (i.e., the left most values are associated with smallholders). Separate trends are given for crop production, livestock production, all on farm production, and on and off farm production where relevant. The bottom row contains two poverty lines for comparability with the per capita real income, where $1.90/capita/day is the poverty threshold for low-income countries and $3.20/capita/day is the poverty threshold for low-middle income countries. Non-parametric trend lines and 95% confidence intervals are given.

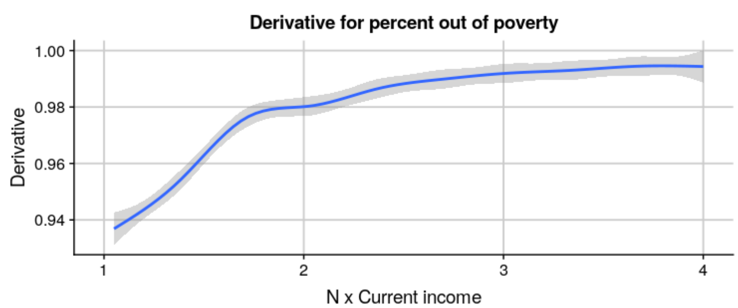

Figure S21: Derivatives for Figure 15A, where for each 0.05 increase in N times the current income there is a given change in the line's slope. At 2x the current income, the slope levels off.

C.2. Supplemental tables

Table S11: Independent variable definitions.

Independent Variable	Definition
Farm size	Farm size (ha); a minimum farm size value of 0.25 ha was assigned due to potential data anomaly introduced in smaller farms.
Relative farm size	Country relative farm size in ranked percentiles.
Relative economic size	Country relative economic size in ranked percentiles of crop revenue.
Female labor	The share of working age (between 15 and 60 years old) household female members to the total number of working age household members.
Family labor	One minus the percent of a household's expenditure on labor compared to total on-farm expenditure, which assumed that family labor was unpaid. We used this definition because it allowed us to measure 30 countries. We tested this definition against the number of days of family labor worked per year over the total number of labor days worked per year per farm for nine countries where we had these variables (this would have been an ideal definition for percent family labor). We observed a very strong linear relationship between our proxy variable and the actual percent family labor (see Figure S19).
Subsistence level	Value of crop consumed divided by the value of crop produced.

Table S12: Fixed effects models predicting per capita income in country relative terms. Coefficients and 95% confidence intervals (in parentheses).

DV	FE1	FE2	FE3	FE4	FE5	FE6	FE7	FE8	FE9	FE10
Farm Size (ha)	0.19*						0		0.15*	
	(0.18; 0.20)						(-0.02; 0.01)		(0.14; 0.16)	
Farm Size (rel)		0.15*						-0.01		0.12*
		(0.14; 0.16)						(-0.02; 0.01)		(0.11; 0.13)
Family Labor			-0.08*				-0.02*	-0.02*	-0.04*	-0.05*
			(-0.10; -0.07)				(-0.03; -0.00)	(-0.03; -0.00)	(-0.06; -0.03)	(-0.06; -0.03)
Subsistence				-0.22*			-0.08*	-0.08*	-0.19*	-0.19*
				(-0.23; -0.21)			(-0.09; -0.06)	(-0.09; -0.06)	(-0.20; -0.18)	(-0.20; -0.18)
Economic Size					0.36*		0.34*	0.34*		
					(0.35; 0.37)		(0.33; 0.35)	(0.33; 0.35)		
Female Labor						-0.06*	-0.06*	-0.06*	-0.06*	-0.06*
						(-0.07; -0.05)	(-0.07; -0.05)	(-0.07; -0.05)	(-0.07; -0.05)	(-0.07; -0.05)
Num. obs.	38270	38270	38270	38270	38270	38270	38270	38270	38270	38270
Num. groups: country:psu	7235	7235	7235	7235	7235	7235	7235	7235	7235	7235
Num. groups: country	30	30	30	30	30	30	30	30	30	30
R^2	0.41	0.41	0.39	0.41	0.47	0.39	0.47	0.47	0.42	0.42
Adj. R^2	0.27	0.27	0.25	0.27	0.35	0.25	0.35	0.35	0.29	0.29

Table S13: Fixed effects models predicting on-farm economic productivity in country relative terms. Coefficients and 95% confidence intervals (in parentheses).

DV	**FE11**	**FE12**	**FE13**	**FE14**	**FE15**	**FE16**	**FE17**	**FE18**	**FE19**	**FE20**
Farm Size (ha)	-0.54* (-0.55; -0.53)						-0.71* (-0.73; -0.70)		-0.55* (-0.56; -0.54)	
Farm Size (rel)		-0.39* (-0.40; -0.38)						-0.52* (-0.53; -0.51)		-0.39* (-0.40; -0.38)
Family Labour			0.13* (0.11; 0.14)				0.09* (0.08; 0.10)	0.11* (0.10; 0.12)	0.06* (0.05; 0.08)	0.08* (0.07; 0.09)
Subsistence				-0.02* (-0.03; -0.01)			0 (-0.01; 0.01)	0 (-0.01; 0.02)	-0.12* (-0.13; -0.10)	-0.11* (-0.12; -0.09)
Economic Size					0.09* (0.08; 0.10)		0.35* (0.34; 0.36)	0.33* (0.32; 0.34)		
Female Labor						0.01* (0.00; 0.02)	0 (-0.01; 0.01)	0.01 (-0.00; 0.01)	0 (-0.01; 0.01)	0 (-0.00; 0.01)
Num. obs.	38114	38114	38114	38114	38114	38114	38114	38114	38114	38114
Num. groups: country:psu	7205	7205	7205	7205	7205	7205	7205	7205	7205	7205
Num. groups: country	30	30	30	30	30	30	30	30	30	30
R^2	0.5	0.48	0.38	0.38	0.38	0.38	0.56	0.53	0.51	0.48
Adj. R^2	0.38	0.35	0.24	0.23	0.24	0.23	0.46	0.43	0.39	0.36