ZENTRUM FÜR ENTWICKLUNGSFORSCHUNG (ZEF)

The role of transaction costs in strengthening agricultural market linkages to achieve higher welfare in Tanzania

Dissertation

zur

Erlangung des akademischen Grades

eines

Doktor der Agrarwissenschaften

(Dr. agr.)

der

Landwirtschaftlichen Fakultät

der

Rheinischen Friedrich-Wilhelms-Universität Bonn

von

André Bueno Rezende de Castro

aus

São Paulo, SP, Brasilien

Bonn, 2023

Angefertigt mit Genehmigung der Landwirtschaftlichen Fakultät der Universität Bonn

Referent: Prof. Dr. Dr. hc. Joachim VON BRAUN

Korreferent: PD Dr. Alisher MIRZABAEV

Tag der mündlichen Prüfung: 17. Oktober 2023

Printed and/or published with the support of the German Academic Exchange Service.

Abstract

Small-scale farmers (SSFs) have high levels of food insecurity and poor nutrition that are reflections of the still high prevalent poverty rates. Semi-subsistence smallholder farming is still predominant in many parts of the developing world and is characterized by severe transaction costs that thwart market access and hinder the improvement of livelihoods. The stark differences in living standards between least developed countries (LDCs) and richer countries are explained by productivity differentials at the frontier, which are pronounced in the agricultural sector. The conceptual framework of this dissertation focused on the role played by high transaction costs in keeping agricultural productivity low and how participation in more sophisticated forms of market access can address this developmental issue. This framework was empirically studied in three analytical chapters exploring different ways whereby improved market access cut transaction costs. The second chapter considered two main groups of vertical market linkages (VMLs) and found positive effects of maize SSFs integration into VMLs on household welfare. Using representative panel data from Tanzania it was found that VMLs centered on agro-processing (AP) firms had larger and more robust effect magnitudes than VMLs centered on cooperatives and informal markets. The third chapter used cross-sectional data collected during fieldwork in Tanzania to explore the connections between SSFs and the AP sector in the form of contract farming (CF): farmers' output and food security increased by 0.13 and 0.4 standard deviations, respectively. Additionally, CF schemes in the cashew nuts sector improve the supply of raw materials to processors thus stimulating an AP-based industrialization. The last analytical chapter shifted the focus from the rural economy to an analysis using cross-country panel data. It studied how the digitalization of customs and border procedures reduces transaction costs, therefore increasing bilateral agrifood trade by at least 9%, particularly in Sub-Saharan Africa and Asia. The main contribution of this dissertation to the fields of agricultural and developmental economics is to bridge the gap of empirical studies analyzing the key role that persistently high transaction costs have on hindering improved market access and, consequently, slowing down rural development. Lastly, and based on the empirical results, this thesis concludes by providing targeted policy recommendations.

Zusammenfassung

Kleinbäuerinnen und -bauern sind in hohem Maße von Ernährungsunsicherheit und Mangelernährung betroffen, was die immer noch hohen Armutsraten widerspiegeln. In vielen Teilen der Entwicklungsländer ist die kleinbäuerliche Semi-Subsistenzlandwirtschaft immer noch vorherrschend. Sie ist durch hohe Transaktionskosten gekennzeichnet, die den Marktzugang erschweren und die Verbesserung der Lebensbedingungen behindern. Die starken Unterschiede im Lebensstandard zwischen Ländern mit niedrigem Einkommen und Industrieländern erklären sich durch Produktivitätsunterschiede an der Grenze, die im Agrarsektor besonders ausgeprägt sind. Der konzeptionelle Rahmen dieser Dissertation konzentrierte sich auf die Rolle von hohen Transaktionskosten, wie sie die landwirtschaftliche Produktivität niedrig halten, und wie die Teilnahme an anspruchsvolleren Formen des Marktzugangs dieses Entwicklungsproblem angehen kann. Drei analytische Kapitel betrachten verschiedene Teile des Rahmens, indem sie eines verbesserten Marktzugangs untersuchen, der Transaktionskosten senkt. Kapitel 2 betrachtete zwei Hauptgruppen vertikaler Marktverknüpfungen (VMLs) und fand positive Auswirkungen der Integration von Mais-Kleinbäuerinnen und -bauern in VMLs auf die Wohlfahrt der Haushalte. Unter Verwendung repräsentativer Paneldaten aus Tansania wurde festgestellt, dass VMLs, die sich auf landwirtschaftliche Verarbeitungsunternehmen (AP) konzentrierten, größere und robustere Effekte hatten als VMLs, die sich auf Genossenschaften und informelle Märkte konzentrierten. Kapitel 3 verwendete Querschnittsdaten, aus Feldarbeit in Tansania, um die Verbindungen zwischen Semisubsistenzbetrieben und dem AP-Sektor in Form von Vertragslandwirtschaft (CF) zu untersuchen: Die Produktion der Landwirte und die Ernährungssicherheit stiegen um 0,13 bzw. 0,4 Standardabweichungen. Darüber hinaus verbessern CF-Programme im Bereich der Cashewnüsse die Rohstoffversorgung des Verarbeitungsunternehmens und ermöglichen so eine agrar-basierte Industrialisierung. Das letzte Analysekapitel verschob den Fokus von der ländlichen Wirtschaft zu einer Analyse unter Verwendung von länderübergreifenden Paneldaten. Es untersuchte, wie die Digitalisierung von Zoll- und Grenzverfahren die

Transaktionskosten senkt und damit den bilateralen Agrarlebensmittelhandel um mindestens 9% steigert, insbesondere in Subsahara-Afrika und Asien. Der Hauptbeitrag dieser Dissertation zu den Feldern der Agrar- und Entwicklungsökonomie besteht darin, die Lücke empirischer Studien zu schließen, die die Schlüsselrolle von Transaktionskosten analysieren, die diese bei der Verhinderung eines verbesserten Marktzugangs und folglich bei der Verlangsamung der ländlichen Entwicklung spielen. Abschließend und basierend auf den empirischen Ergebnissen liefert diese Thesis gezielte politische Empfehlungen.

Acknowledgements

First and foremost, I thank my supervisor, Prof. Dr. Dr. h.c. Joachim von Braun, for having accepted me as a doctoral student as well as for his kind support throughout my doctoral journey and insightfully provocative comments and suggestions. Second, I thank PD Dr. Alisher Mirzabaev for having accepted to be my second supervisor as well as Prof. Dr. Jan Börner and Prof. Dr. Christian Borgemeister for having accepted to be member of my defense committee. I also thank and acknowledge the German Academic Exchange Service (DAAD) for funding my doctoral studies with both a scholarship and research funding, and, likewise, I gratefully acknowledge the Dr. Hermann Eiselen Foundation fiat panis for funding part of my fieldwork budget. Also related to my fieldwork stay in Tanzania, I further acknowledge the School of Economics of the University of Dar es Salaam – especially Prof. Dr. Beatrice Mkenda, Dr. Martin Chegere, and Dr. Onesmo Selejio – and the Research on Poverty Alleviation (REPOA) institute – especially Dr. Jamal Msami and Dr. Donald Mmari – for having hosted me. I also thank all Government of Tanzania officials who have received and helped me during my fieldwork, especially the Cashewnuts Board of Tanzania; I sincerely hope that my studies can be useful for the country's policymakers. Last but not least, I gratefully thank the farmers and firm managers I interviewed and my research assistants and enumerators, Dr. Kevin Rugaimukamu, Farida Katunzi, Daniel Sifael, and Peter Nyanda.

I appreciate very much all the support and guidance I received from my academic tutor, Dr. Lukas Kornher, without whom this dissertation would not have been possible. The same gratitude extends to other ZEF senior researchers who have helped me better shape my dissertation, especially Prof. Dr. Matin Qaim, Dr. Zaneta Kubik, Dr. Jorge Luis Sellare, and Dr. Gustavo Magalhães de Oliveira, as well as colleagues from the Volatility Team. I also appreciate the friendship and mutual support from all ZEF junior researchers, especially, but not only, from my BIGS-DR batch 2019 as well as all ZEF staff, in particular Dr. Silke Tönsjost, Maximilian Voit, Volker Merx, and Anja Köhler. All three analytical chapters from this dissertation were presented in workshops and conferences and I thank all participants who commented and helped me improve the quality of my work. I extend the same acknowledgment to all journal editors and reviewers who revised my submitted manuscripts. I am also specifically thankful for comments from friends and colleagues as well as discussions about either my individual chapters and/or the dissertation: Prof. Dr. Jann Lay, Dr. Annekathrin Schoofs, Dr. Sahrah Fischer, Dr. Mekdim Dereje Regassa, Dr. Isaac Bonuedi, Dr. Henry Kankwamba, Dr. Shaibu Mellon Bedi, Yauheniya Shershunovich, Hina Amber, Paulo Mortara Batistic, Jedediah Silver, and Kwabena Nyarko Addai.

Finally, I express my deepest gratitude to all my friends and family, especially my parents, for their care and unconditional support. Antonia Ortmann, I will be forever grateful to you for your patience and companionship. I am very much looking forward to our next adventures. Alina Ortmann Castro, you are the joy of my life and I am very happy you could witness my defense. I will always be there for you.

“[…]

As coisas tangíveis

tornam-se insensíveis

à palma da mão.

Mas as coisas findas,

muito mais que lindas,

essas ficarão.”

“Memória”, Carlos Drummond de Andrade

Table of Contents

List of Tables

List of Figures

List of Abbreviations

AfCFTA	African Continental Free Trade Area
AfDB	African Development Bank
AMCOS	Agricultural and Marketing Cooperative Society
AP	Agro-processing
AME	Average marginal effect
ANSAF	Agriculture Non-State Actors Forum of Tanzania
ATE	Average treatment effect
ATT	Average treatment effect on the treated
AU	African Union
BACI	International Trade Database at the Product-Level
CB	Paperless trade for cross-border NTMs
CBT	Cashewnuts Board of Tanzania
CEPII	Center for Prospective Studies and International Information
CF	Contract farming
COVID-19	Coronavirus Disease 2019
CTI	Confederation of Tanzania Industries
CRE	Correlated Random Effects Model
DAAD	German Academic Exchange Service
EU	European Union
FAO	Food and Agriculture Organization of the United Nations
FD	First Differences Model
FDI	Foreign direct investment
FE	Fixed Effects Model
FEGLS	Fixed effects generalized least squares
FYDP	Five-Year Development Plan
GDP	Gross domestic product
GoT	Government of Tanzania
GVC	Global value chains

HDDS	Household Dietary Diversity Score
HDFE	High-dimensional fixed effects
HFIAS	Household Food Insecurity Access Scale
HIC	High-income country
HM	Household member
ICT	Information and communication technologies
IPWRA	Inverse Probability Weighted Regression Adjustment
LDC	Least developed country
LMIC	Low- and middle-income country
MoA	Ministry of Agriculture of Tanzania
NBS	National Bureau of Statistics of Tanzania
NDV25	National Development Vision 2025
NPS-LSMS	National Panel Survey-Living Standards Measurement Study
NTM	Non-tariff measure
OLS	Ordinary Least Squares
PPML	Pseudo-Poisson Maximum Likelihood
PSM	Propensity Score Matching
PT	Paperless trade
RCA	Revealed Comparative Advantage
RCN	Raw cashew nuts
RE	Random effects model
RTA	Regional trade agreement
RVC	Regional value chains
SDG	Sustainable Development Goals
SITC	Standard International Trade Classification
SME	Small and medium enterprises
SPS	Sanitary and phytosanitary standards
SSA	Sub-Saharan Africa
SSF	Small-scale farmer
TFA	Trade facilitation agreement

Tsh	Tanzanian shillings
UN	United Nations
UNTFS	United Nations Global Survey on Digital and Sustainable Trade Facilitation
USD	United States Dollar
VML	Vertical market linkage
WTO	World Trade Organization
ZEF	Zentrum für Entwicklungsforschung / Center for Development Research

CHAPTER 1

Introduction

1.1 Background and motivation of the study

In the vast majority of countries in Sub-Saharan African (SSA) agricultural production is conducted in small plots of land by small-scale farmers (SSFs)[1], who have low degrees of crop commercialization, thus relying on semi-subsistence[2] activities (Lowder et al., 2016). Poor market access is particularly prevalent among SSFs and it poses a major constraint for better rural welfare (Frelat et al., 2016). The reason is that greater market access lowers transaction costs by reducing input and transportation costs as well as increasing output prices, hence leading to higher agricultural incomes (Dercon et al., 2009; Stifel & Minten, 2017). The main objective of this thesis is to investigate the role that transaction costs play for SSFs agricultural commercialization and connection to agricultural market linkages and, consequently, how it impacts rural welfare. Additionally, I explore how improved connections to such market linkages are important steps in the process of transformation of agrifood systems in least developed countries (LDCs). Specifically, I hypothesize that the reduction of transaction costs associated with a connection to more sophisticated agricultural market linkages improves SSFs' incomes and food security thus leading to higher rural welfare. This hypothesis is derived from a harmonized conceptual framework discussed in Section 1.4 and is tested empirically in different contexts in the three analytical chapters. In other words, the main research question that this dissertation answers is as follows:

- Do agricultural market linkages reduce transaction costs for SSFs and increase rural

[1] I adopt the term "small-scale" rather than "smallholder" farmers to also include farmers that do not produce on large scale in terms of total output. "Smallholder" characterizes farmers by directly connecting them with the land's plot size while "small-scale" also encompasses production scale and input utilization intensity.

[2] Subsistence agriculture means that a farmer or a farming household derives all of their food consumption from what they produce, so the term semi-subsistence is preferred to capture the spectrum ranging from full subsistence to full commercialization.

welfare? What are the possible transmission mechanisms?

The importance of agriculture in SSA economies – and in LDCs in general – can be perceived by the large share of value added as a percentage of gross domestic product (GDP): in 2019 it accounted for 16.5%, and the employment share was 53% (World Bank, 2020b)[3]. There are three important remarks about these twin figures: (1) the share of agricultural value added has been stagnant for about four decades whilst the agricultural employment share has slightly decreased; (2) both are still high shares relative to other developing countries, which indicates a protracted process of structural transformation (Lewis, 1954); (3) it is noticeable that the employment share is still well above the agricultural share of GDP, which reflects the fact that employment in agriculture is often part-time, but also suggests that the agricultural labor productivity gap within the country is yet to be closed (Gollin et al., 2014b). Agriculture and the agro-processing (AP) industry combined[4] are the backbone of SSA economies and play a fundamental role for rural livelihoods by ensuring food security, earning foreign currency in exports, and generating income at the local level. This combined sector is also essential to generate backward and forward linkages to other economic sectors. Notwithstanding the key importance of agriculture in LDCs, support to this sector is often overlooked by policymakers – mostly, but not only, in terms of budgetary priorities (World Bank, 2008).

Naturally, agricultural production is not fully subsistence as approximately 80% of SSFs in SSA sell some share of their agricultural output even if they do not grow enough crops to become self-sufficient (Wiggins & Keats, 2013). However, the agriculture's commercialization component largely aims at supplying domestic and local markets with weak connections to global value chains (GVCs) following colonial patterns of low value addition in exports of raw materials (Roessler et al., 2022). Over three-quarters of African households are involved in

[3] These figures are even higher in Tanzania: 26.7% and 65%, respectively.
[4] I avoid using the term "agribusiness" as a general description of the main object of analysis of this dissertation because it often refers to more commercialized, capitalized, large-scale, and, often, vertically integrated forms of agricultural and agro-processing production, thus excluding the majority of the rural activities in the country.

some way (part-time or full-time) in agricultural activities, either as subsistence farmers, selling crops to the markets, or as farm laborers (FAO et al., 2022). The majority of agrifood products in SSA (markedly in LDCs) still comes from semi-subsistence farming, with a slowly growing share from small and medium commercially-oriented family farms. The defining characteristic of subsistence farming are high transaction costs and low productivity, in terms of yield per farmed area (Gollin & Rogerson, 2014). Difficulties in increasing agricultural productivity are associated with persistently high transaction costs that block SSFs' market participation (de Brauw & Bulte, 2021), which is an issue analyzed in this dissertation. One major drawback of predominant semi-subsistence farming is that it often leads into a low-level equilibrium, i.e., a poverty trap, from which a sustainable escape is achieved either via a cash windfall (i.e., from government programs) or via technology transfers (Manh Hung & Makdissi, 2004); which is hard to achieve for credit constrained economies[5]. In contrast, crop commercialization is associated with gains from specialization and technology adoption, which increase crop yields and incomes (Bellemare et al., 2022).

Such high reliance that a large share of livelihoods has on low-productivity semi-subsistence agriculture has negative developmental consequences. The prevalence of severe food insecurity and of severe or moderate food insecurity in the past 5 years rose by 5 percentage points (p.p.) to 24.8% of the SSA population and by 10 p.p. to 60.9%, respectively (FAO et al., 2022)[6]. The corresponding figures for Tanzania are similar increases of 5 p.p. and 9 p.p., respectively, with respective figures of 25.8% and 57.6% of the country's population, which puts it slightly better-off than the sub-continental average. Nevertheless, some improvement has been achieved in the prevalence of undernourishment, which decreased by almost 6 p.p. to 22.6% in the last 15 years in Tanzania, compared to a smaller reduction of 2.4 p.p. to 22% in SSA.

[5] LDCs often suffer from expensive credit or lack thereof.
[6] This FAO data does not include the negative global effects on food, fuel, and agricultural inputs prices caused by the Russian invasion of Ukraine and the ensuing food crisis, so these figures will likely increase.

Another trend that can further strain developmental and livelihood improvements is the persistently high population growth rate. The population in LDCs is forecast to continue expanding at the world's highest population growth rates, in particular in SSA, which is projected to double its current 1.18 billion inhabitants by 2055 (UN, 2022). Such high population growth can contribute in the opposite direction of recent improvements in the reduction of undernourishment[7] (FAO et al., 2022). Although this phenomenon is not analyzed per se in this dissertation, it carries two main implications that are relevant to the analytical chapters. First, the higher the population growth, the more the labor market is pressured to generate a higher quantity of jobs. This is especially true for SSA where the recent economic growth was not able to absorb labor at a quantity sufficient to improve livelihoods (AfDB, 2019; Kubik et al., 2022). Second, and as a consequence of the first implication, the consumption of food is projected to significantly grow, not exclusively because of a larger population that needs to be fed, but also significantly because of an ongoing shift in dietary patterns in low- and medium-income countries (LMICs). According to Bennett's law (Bennett, 1941), following sustained periods of economic growth consumers from LMICs gradually modify their food baskets by substituting calories from carbohydrates derived from starchy staples (i.e., maize, cassava, sorghum) and coarse grains in benefit of finer grains (especially rice and wheat), fruits, vegetables, livestock products, and processed food products (Fukase & Martin, 2020; Gouel & Guimbard, 2019). This is a pattern observed both in SSA in general (Reardon et al., 2021; Tschirley et al., 2015) and in Tanzania in particular (Sauer et al., 2021). What is highlighted from these phenomena is the pressing need to increase agricultural production through the intensive margin – higher agricultural productivity – and not through the extensive margin – adding factors of production.

[7] For example, while the prevalence of undernourishment fell, the number of undernourished people increased by by 63.2 million and 2.6 million in the same 15-year period in SSA and in Tanzania, respectively. Similarly, while the prevalence of stunting in children under 5 years of age fell, the number of stunted children remained almost the same.

The current state of the agricultural sector in SSA is under its potential (Barrett et al., 2017; Gollin & Udry, 2021; Kirui et al., 2023). Key to unlock this potential are joint and coordinated efforts from the public and the private sectors to connect SSFs to agricultural market linkages that can cut transaction costs, thus harnessing the structural transformation process (Barrett et al., 2017). This is often difficult because there are substantial gaps between agricultural and non-agricultural productivity in developing countries (Gollin et al., 2014b), thus the systematic shift of productive factors out of agriculture into more productive sectors (i.e., structural change/transformation) has meaningful implications for aggregate labor productivity and, consequently, for living standards (Adamopoulos & Restuccia, 2022). For instance, the highest share of the economy-wide labor productivity growth in SSA comes exactly from the part attributed by Diao et al. (2019) to structural change, while the rest comes from agricultural labor productivity growth per se. Two implications emerge from this stylized fact: first, agricultural labor productivity growth is still small considering the heretofore labor productivity gap with respect to modern sectors and considering the relatively stagnant crop yields[8]; second, very little of SSA's labor productivity growth is coming from technological change, which does not bode well for a sustained growth rate (Diao et al., 2019; McMillan & Zeufack, 2022). Ultimately, stronger growth in agricultural productivity – a topic addressed in the analytical chapters – boosts commercialization, which has long been seen as a key lever to lift the rural economy from poverty via higher yields and incomes (von Braun, 1995). Notwithstanding, an important caveat about the potential of commercialization is due:

> Market participation is a consequence as much as a cause of development. Just ‘‘getting prices right” does not induce broad-based, welfare-enhancing market participation. Farm households must have access to productive technologies and adequate private and public goods in order to produce a marketable surplus. One thus has to get institutions and endowments, as well as prices, ‘‘right” in

[8] Yields disaggregated by crops can be found here: https://ourworldindata.org/crop-yields <Accessed on 10 November 2023>.

order to induce market-based development. (Barrett, 2008, p. 300).

The main agricultural production in Tanzania features staple foods (particularly maize) – with a growing, but still small, share of fruits and vegetables (NBS, 2021). The cultivated area is divided into approximately one million agricultural units, of which around 80% are in the hands of smallholder farmers[9] with plots[10] averaging 1.5 hectares (NBS, 2021). Tanzania is a useful study region for the research questions of this dissertation because of two main reasons: (1) it is an important agricultural player in the continent[11]; (2) it has characteristics that allow generalization of the findings, to some extent, to other similar LDCs, particularly in SSA. Some of these generalizable characteristics are: high proportion of the workforce in agriculture, i.e., ongoing structural transformation; relatively high combined GDP share of agriculture and AP industry; tropical and semi-tropical agroecological zones; export dependence on a few agricultural and mineral commodities; subordinate insertion into the GVCs underpinned by enduring power asymmetries; the recent colonial history that still reflects in current socioeconomic conditions and institutions (Roessler et al., 2022).

Furthermore, Tanzania is a large and rapidly growing LDC that has undergone a significant amount of growth volatility and varying types of industrial policy regimes since gaining independence in 1961. This is due to the fact that Tanzania has moved to a market economy relatively late compared to other developing countries; in the 1980s-1990s it started the process of switching from a model of trade protection, price controls, and import substitution to another

[9] The term smallholder farmer has diverse meanings both in the literature and between countries and contexts. The IFC (2013) defines smallholder farmer as a family-owned agricultural enterprise that grows crops and/or livestock in landholdings of 2 or less hectares. Given the focus on Tanzania and its land structure, this definition fits fairly well, although it might exclude some farms that are just marginally above the 2-hectare threshold. Therefore, for the purposes of this thesis, I consider a smallholder farmer as a farm that is owned and/or tilled by a family and that has little mechanization and reliance on hired labor. Furthermore, in order to capture nuances in agricultural production that are beyond the size of the plot of land, the term small-scale farmer is preferred to smallholder – see footnote 1.
[10] A household can own and/or cultivate more than one plot of land (or parcels of land).
[11] Tanzania is a leading exporter of tobacco, cashew nuts, coffee, maize, and sisal. Data from FAOSTAT: https://www.fao.org/faostat/en/#data/QCL <Accessed on 10 November 2023>. Tanzania has about 5% of the combined African agricultural value added. Data from the World Bank's World Development Indicators: https://databank.worldbank.org/source/world-development-indicators <Accessed on 10 November 2023>.

model based on macroeconomic liberalization with export promotion and uncoordinated active industrial policy mechanisms (McMillan et al., 2017; Msami & Wangwe, 2016). Moreover, Tanzania is geolocated in a strategic trade hub in East Africa and is the regional leader in the trade of some staple grain crops, such as maize and rice[12].

Finally, there are many crops for which the country has adequate agroecological conditions to be a powerhouse, which means not only supplying the domestic demand, but also exporting large surpluses to the world market in exchange for scarce foreign currency. Based in order of importance in the Tanzanian export basket, the main examples are: tobacco, cashew nuts, coffee, dried legumes, oily seeds, fish, maize, and tea. The potential of cashew nuts is further explored in Chapter 3, being a crop that has been extensively studied in West Africa but for which there is a gap of empirical studies focused on Tanzania (Danso-Abbeam et al., 2022; Dubbert, 2019; Miassi et al., 2019). Other studies, especially in East Africa, have already explored the potential of fruits and vegetables both as an export cash crop and as a lever to transform rural areas (Ajwang, 2020; Benali et al., 2018; Mergenthaler et al., 2009; Rao et al., 2012). However, there is a lack of studies on the potential of food crops, such as maize – explored in Chapter 2 –, to help SSFs to connect to agricultural market linkages (Biggeri et al., 2018; Herrmann et al., 2018; Mmbando et al., 2015; Montalbano et al., 2018). And Chapter 4 considers the aggregate of agrifood production using cross-country trade data.

1.2 Research objectives and contributions to the literature

The overarching research objective of this dissertation is to investigate the effects that cutting transaction costs related to access to markets and participation in improved market linkages have on rural household welfare and agrifood trade. This dissertation builds on the literature on agrifood value chains/systems in LDCs (Barrett et al., 2018, 2022) and goes beyond theoretical

[12] Data from UN Comtrade: https://comtrade.un.org/ <Accessed on 10 November 2023>.

studies by providing empirical evidence on how SSFs can connect to agricultural market linkages and its effects on their welfare (Bellemare et al., 2022). This main research objective is sub-divided into three specific research questions that are answered in three analytical chapters which contain empirical studies that vary in scope and meaning but are tied into the same storyline. Each chapter contributes to the agricultural and development economics literature through a different aspect of SSF's agricultural market integration, also discussing possible transmission mechanisms for the effects identified. The materials and methods employed in the analytical chapters are adaptable and can be applied to a wide variety of countries and regions as well as different research questions. Specifically, two of these analytical studies have a country focus on Tanzania and one of them is a cross-country case. The specific research questions are as follows:

1. What are the effects on household welfare and agricultural production of Tanzanian SSFs' participation in vertical market linkages (VMLs)?
2. What are the effects on food security and cashew nuts production of Tanzanian SSF' integration with AP firms through contract farming (CF) schemes?
3. What are the level and effects of transaction costs related to trade and customs digitalization of procedures and bureaucracies on the diversification of SSA agrifood trade?

The scope of this dissertation provides empirical evidence related to the sustainable development goals (SDGs) #1 (no poverty), #2 (zero hunger), #8 (decent work and economic growth), and #9 (build resilient infrastructure, promote sustainable industrialization, and foster innovation). Moreover, in terms of policy outreach, the body of the work of this dissertation speaks directly to governmental priorities as industrialization is of paramount importance for the Government of Tanzania (GoT) and plays a key role in the last three Five-Year Development Plans (FYDP). The last FYDP refers to the financial years 2021/22-2025/26 – all integrated into the Long-Term Perspective Plan 2011/12-2025/26. In practical terms, the FYDPs are the roadmap for the GoT to achieve the goals set in the National Development Vision 2025 (NDV25) of improving the standards of living towards a lower middle-income status. The NDV25 overarching goal aims to utilize the available natural resources to increase the economy's efficiency and productivity, jumpstarting an industrialization process. Thus, the direct connection between the latest FYDP and the NDV25 with the main topic of this dissertation allows this work to propose targeted recommendations in the concluding chapter.

1.3 Definitions of concepts

This sub-section explains how the concept of transaction cost is employed in the analytical chapters, with an extension to agriculture, thus providing a theoretical basis for the conceptual framework. Also defined here are two other important concepts: agrifood value chains and agro-processing industry.

First, it is important to distinguish transaction costs from ordinary production costs. The former can be defined more broadly as the cost of economic coordination when using the price system (Coase, 1937) or, more narrowly, it is the search and matching costs involved with the transfer of property of a certain commodity between two or more economic actors (Niehans, 1987). These search costs involve not only costs related to searching for exchange partners but also screening costs (gathering of information about an exchange partner's reliability and quality of goods/services) and contract monitoring and enforcement costs (Stigler, 1972). According to the definition of transaction costs used in this dissertation, transaction costs do not include direct agricultural production costs, such as labor, land, capital, and other inputs, rather they include the organizational and coordination costs of marketing/selling the product along the supply chain up to the final consumer. In a sense, it could be seen as the equivalent to the role of friction in physics (Stigler, 1972). This dissertation directly borrows the concept of transaction cost from the New Institutional Economics school of thought (Williamson, 1979) to explain why the shift from semi-subsistence to a more commercially-oriented agriculture takes place slowly, despite its crucial role in the ongoing process of structural transformation (Dorward, 2001; Gollin et al., 2002).

Transaction costs also play an important role in trade, somewhat similar to the iceberg model of transport costs. For instance, digitalization measures can improve access to agricultural markets, reduce price dispersion, and increase the efficiency of market coordination (Aker, 2010; Jensen, 2007; Jensen, 2010). Although these studies focus on the initial impact of the introduction of mobile telephones in developing countries, the frameworks used allow us to consider their findings for a wider definition of digitalization. In a similar manner, digitalizing customs and trade procedures, as analyzed in Chapter 4, is a significant step forward in reducing the toll that such transaction costs exert on trade partners, thereby increasing trade.

In the context of agriculture, the importance of transaction costs is reflected in the market structure and coordination mechanisms. These costs are inherently connected with market failures, which are abundant in LDCs, such as uninsurable risks, missing markets and contracts, information asymmetry, and ill-defined property rights (Morrissy, 1974). Transaction costs in agriculture in LDCs can be so high that they either thwart market transactions, resulting in inefficiencies, or prevent the formation of formal inputs and outputs markets[13] altogether, i.e., the margins between sales and purchase prices become so wide that markets fail (Janvry et al., 1991; Key et al., 2000). Consequently, when the cost of transactions through market exchanges exceeds the utility and/or the revenues accrued by them, the producer would likely refrain from utilizing such market system, ultimately reverting to informal market arrangements or even to subsistence. This conclusion bears resemblance to the theory of the firm (Coase, 1937) with one difference: the unit of analysis is not the firm but the farming household. This framework

[13] A key difference between formal and informal markets is the probability that market exchanges are regulated and laws enforced by a third party, frequently the government (de Brauw & Bulte, 2021).

defines if economic activities occur within the market or within the firm, in which case the magnitude of transaction costs precludes a market arrangement (Williamson, 2002), as noted above. This is one of the main explanations for the pervasiveness of semi-subsistence in rural SSA vis-à-vis the still relatively low levels of agricultural commercialization (Janvry et al., 1991). Greater participation in agricultural markets, especially when they are formalized, would, in turn, greatly increase agricultural productivity and contribute to rural poverty alleviation (von Braun, 1995).

For non-cash crops, SSFs' agricultural marketing in LDCs is often not structured and is associated with significant transaction costs as noted above. This problem also affects cash crops and one solution that SSA countries adopted, just after independence, was to use parastatal marketing boards to channel their cash crop exports. One example is the Ghanaian cocoa board that coordinates all raw cocoa sales and plays a key role in facilitating exports of processed beans and financing for farmers and local processors (Huellen & Abubakar, 2021). Another example is the Tanzanian cashew board, with analogous functions although with more limited capabilities, that is analyzed in the third chapter. Moreover, this definition of transaction cost is also informative to understand the challenges associated with CF and VMLs analyzed in Chapters 2 and 3: *ex-ante* information costs of discovering prices and screening for buyers/sellers, contract negotiation and signing costs, and *ex-post* costs of contract monitoring and enforcement (formal or informal) – as noted above. In addition to parastatal marketing boards, these two analytical chapters study alternative market coordination mechanisms, embedded in the framework of improved smallholder access to markets through agricultural marketing linkages. Government-run auctions of raw cashews and CF schemes have the objective of reducing information asymmetries, making access to finance less of a binding constraint, and coordinating the supply of raw cashews with the demand of local processors.

The second concept defined here is agrifood value chain[14] (often referred to as 'agrifood system'). It is broader in scope than primary agricultural production and food processing because the term also encompasses production factors (land, labor, and capital), agricultural input suppliers (such as enterprises producing seeds, fertilizers, and farm machinery), and companies that distribute and sell products to final consumers (such as wholesalers, food retailers, and food service companies). The term 'agrifood value chain' refers to all economic activities performed by a collection of farms, industries, and markets (both formal and informal) that interact to produce, add value at each stage, and distribute food and other agricultural products to AP firms and to the final consumer (Gollin et al., 2002; Kaplinsky & Morris, 2000) – in sum, the combination of the producers upstream and the processors downstream (Bellemare et al., 2022).

Lastly, agriculture is the first link in the supply chain that provides food for consumers and raw materials for the AP industry. This supply chain stretches far beyond the farm gate through several processes of value addition to raw materials and also provides inputs to agriculture itself. The term 'agro-processing' refers to this whole complex of industrial production processes as well as associated services that use raw materials from the agrifood sector, chiefly among these are storage, transport, processing per se, and the wholesale and retail of final products (Tschirley et al., 2015). It is then the industrial sector that processes food and other agricultural raw materials through mechanical, biological, and chemical processes to serve either as inputs for further industrial processing or as an end product for consumption (Hollinger & Staatz, 2015). The AP sector is, thus, the one with the closest interlinkages between agriculture and industry.

[14] The term 'agrifood' is used instead of only 'agriculture' or 'food' because 'agrifood' encompasses the entirety of the products and sub-products from the primary sector.

1.4 Conceptual framework

This sub-section is anchored on two related processes: first, the stages of transformation of agrifood value chains; and second, structural transformation of the economy. These two processes are inherently endogenous to the dynamics of an economy and, as such, will not be a direct object of analysis of this dissertation but rather serve as a theoretical basis for the conceptual framework. I then analyze the connections between the stakeholders shown in the framework.

Following the definitions by Barrett et al. (2022) and Reardon et al. (2021), the majority of LDCs in SSA are moving from the traditional stage – more predominant in rural areas – to the transitional stage – more predominant in large urban areas – of its agrifood sector. This means that the agrifood value chain is becoming spatially wider and longer in terms of the number of links, i.e., there are more complex activities taking place. This has the following direct implications: (1) the market structure is still fragmented between a large number of small and medium enterprises (SMEs) and a handful of dominant large enterprises; (2) the most used technologies are still labor-intensive due to this factor's relative abundance and credit constraints to purchase capital, but there are some signs of increasing use of capital-intensive technologies; (3) slow shift from spot market transactions towards contracts and/or repeated relationships (key aspects in this dissertation analyzed in Chapters 2 and 3); (4) emergence of simple public health and quality standards for agrifood products (discussed in Chapter 4). The predominance of formal contracts in economic relationships (not only CF but also in all other activities in the value chain) and the ability of the government to enforce these contracts are key pre-conditions for the agrifood value chain to transition from the traditional to the complex stage (de Brauw & Bulte, 2021). This poses a major challenge for policymakers that have to devise public policies to promote this transitional process.

The transformation process of agrifood value chains takes place within the context of the process of structural transformation; both combined explain the main connection of the conceptual framework, i.e., between SSFs and AP firms. In this process, the rural economy is linked with the urban economy through excess low-skilled labor that migrates to larger cities and/or off-farm employment in peri-urban areas or secondary towns seeking better work conditions and/or income opportunities (Christiaensen et al., 2013; Mellor, 2017). This happens because agricultural productivity growth, that tends to go hand in hand with economic growth in LDCs, leads to a reallocation of agricultural labor to more productive sectors, initially manufacturing, then more complex services[15], whose wages subsequently grow (Gollin et al., 2002). This phenomenon is especially pronounced in LDCs in SSA, which usually have low agricultural productivity, thus requiring a large workforce in the fields to produce a given farm/forestry output (Henderson et al., 2018).

Capital accumulation usually starts in the agriculture sector with the profits being diverted to fund industrial and trade activities. The literature indicates that productivity gains in agriculture are fundamental drivers of this transfer of more qualified labor between sectors, thus promoting an economy-wide productivity increase via structural transformation (Adamopoulos & Restuccia, 2022; Gollin et al., 2014b; Timmer, 2002). There are three main mechanisms whereby improvements in agricultural productivity impact economic development, broadly speaking (Mellor, 2017): (1) it directly increases the income of rural producers, a topic addressed in the second and third chapters; (2) it increases the supply of food, so efficiency gains reduce domestic prices – and, consequently, living costs – especially for the poorest, who spend a higher proportion of their budget on food (Fukase & Martin, 2020); (3) it enables meeting the food needs of the population with a smaller number of workers. This releases labor for employment in more productive non-agricultural activities that are essential to promote

[15] There is a trend in LDCs, particularly in SSA, of low-skilled labor shifting from agriculture directly to low-productivity non-tradeable services (such as petty trade, informal transportation, and customer personal services) instead of to manufacturing, a dynamic sector that is more formal and offers higher wages and better working conditions (McMillan et al., 2014; Rodrik, 2016).

broader economic development, especially export generating ones (Lagakos & Waugh, 2013), a topic addressed in the third and fourth chapters.

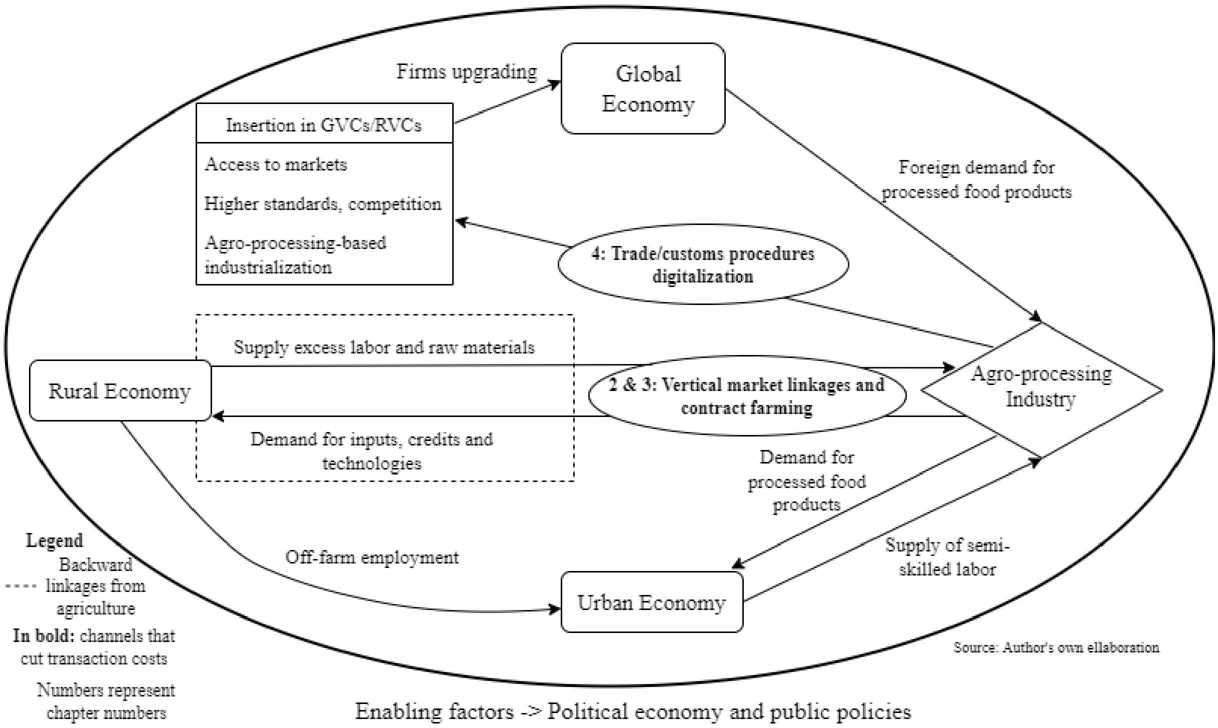


Figure 1.1: Conceptual framework of the dissertation highlighting four main axes (AP firms, rural economy, global economy, urban economy)

The conceptual framework of this dissertation places the main research questions and objectives of the analytical chapters into a coherent storyline. It should not be interpreted as claiming causalities and follows closely a GVC framework (Gereffi et al., 2001), wherein the flow of products and inputs along the value chain is identified, thus allowing to scrutinize the roles of each actor as well as the distribution of gains and where the opportunities for upgrading and value addition lie. Figure 1.1 contains four main axes from which the analytical chapters analyze in depth the first three: AP firms, rural economy, global economy, and urban economy. The numbers of the analytical chapters are highlighted in bold and reflect different channels whereby transaction costs can be cut favoring market access, market integration, and/or international trade. The main connection analyzed in this dissertation is between AP firms and SSFs and is depicted across the main horizontal axis of Figure 1.1. In this connection, farmers supply firms with raw materials for processing and also with labor (including off-farm part-time jobs), whereas farmers demand inputs, credit, and technologies.

Moving to the bottom right of the conceptual framework, AP firms then connect with the urban economy directly by providing food products and demanding labor, frequently low- or semi-skilled. This relationship is thwarted because SSA countries are only slowly reducing their share of workers in agriculture, partly a consequence of slow agricultural labor productivity growth in the last three decades (Masters et al., 2018). Last, the top of Figure 1.1 captures the connections of the stakeholders with the global economy, since chapters 3 and 4 deal with an open economy. International markets demand food products with varied degrees of processing – depending on the destination market – from local AP firms. This is a potential channel for an industrialization push that is based on an export-led growth strategy expansion of exports from the AP sector (Estmann et al., 2022). However, this depends on a more qualified insertion of firms and farmers into the GVCs, including functional and product upgrading. Two important enablers of this insertion are improved access to markets (as addressed in Chapters 2 and 3) and higher products standards to increase competitiveness (as discussed in Chapter 4). International trade has the potential to be a conduit to stimulate an AP-based industrialization in Tanzania because it is a small low-income economy and, as such, it has a limited domestic market.

This framework is relatively simple in terms of the connections and interactions between the actors. Its simplicity is in itself useful to make the analysis of the research questions from Section 1.2 clear as well as the hypotheses embedded in the research questions testable; nevertheless, some real-world features cannot be endogenously explained in the framework. One key concept that is captured in the framework and is included in the analyses in each analytical chapter is transaction cost – the definition is discussed in Section 1.3. In this sense, there are some important conditions for the interconnections and linkages in this framework to happen efficiently: improve access of SSFs to output markets, integrate suppliers and processors into integrated production systems (some of these CF schemes), and advance the level of information and communication technologies (ICTs) and digitalization of trade and customs procedures to improve international competitiveness.

The last part of the conceptual framework is overarching across all its elements and is at the bottom of Figure 1.1. It is related to political economy aspects and encompasses two important enabling factors related to governmental action and the relationship between the government and the private sector. The effective implementation of public policies is often hampered by two shortcomings: (1) weak governance and planning, which is pervasive across all levels of the GoT, with some pockets of excellence; (2) political capture, which happens when private actors accumulate a certain high level of wealth (or attain a high market share) and control over income flows enabling them to acquire political power and influence[16] policy making.

1.5 Outline of the thesis and preview of findings

After this introductory chapter, three analytical chapters (Chapters 2, 3, and 4) follow containing original empirical studies, that provide answers to the research questions presented in Section 1.2. Each one of them has its own introduction and conclusion, thus standing alone as individual papers while simultaneously being components of the main story line. Chapter 5 closes this dissertation with a general conclusion that connects findings as well as the most important policy implications and recommendations arising from the analytical chapters, in addition to some directions for future studies based on the limitations of the analyses. Lastly, the appendices contain extra tables and figures to complement the analytical chapters.

Chapter 2, the first analytical chapter, showed that market integration through agricultural value chains is an important goal for lifting SSFs out of poverty. Focusing on Tanzania, this chapter investigated how the integration of SSFs into VMLs increases household welfare. The contribution to the literature specialized on agricultural value chains was done in two ways. First, I considered a wider variety of welfare indicators (i.e., income, poverty, food security,

[16] The high degree of market concentration in many LDCs incentivizes a few large firms to seek individual lobby with the government rather than constituting effective business associations, leading to an unbalanced amount of influence that such firms exert over the government.

and subjective well-being) in addition to crop yields and commercialization. Second, this chapter analyzed the heterogeneous welfare effects of two different channels: one channel is composed by connection to farmers' cooperatives, local vendors, and traders; the other considers connections to AP firms and to other more formalized businesses as well as CF. The hypothesis that VML participation cuts transaction costs and, thereby, increases SSFs welfare is tested using representative countrywide balanced panel data from Tanzania (NPS-LSMS). Specifically, the methodology employed fixed effects and the first-difference estimators to control for unobserved household time-invariant characteristics; propensity score matching (PSM) is used to address SSFs' selection bias from observed factors. The most important finding of this chapter is that integration into VMLs increases household welfare for all outcome variables in the processor VML group, but less so for the cooperatives group; for income the effect is strong and positive in both groups. However, direct production and income gains do not always convert into enhanced welfare, especially in terms of food security. This means that the issue of which specific pathways enable agricultural value chains to effectively benefit SSFs in LDCs, such as Tanzania, should receive additional attention from future research.

Chapter 3 keeps the focus at the country-level on Tanzania but now considering a specific cash crop: cashew nuts. It has had prominence in the country for decades, bringing hard currency, and being the mainstay of the majority of farmers' households. CF is an institutional innovation that has been extensively studied in a wide variety of contexts, however there is a literature gap on its implications to the cashew nuts value chain development, especially in the case of Tanzania. CF has been recently introduced in the country and, although it holds the promise to improve the sector by directly connecting farmers with AP firms, it still has challenges that need to be addressed with more targeted policies. Within this context, this study fills the literature gap on the welfare effects of CF in the Tanzanian cashew nuts sector by conducting an impact assessment of the effects of marketing CF schemes on tree cash crop production and food security. The methodology utilized an instrumental variable (IV) approach to control for reverse causality and unobserved confounders, and the control variables were added to the preferred specification in a stepwise manner. PSM and the inverse-probability weighted regression adjustment (IPWRA) were employed as further robustness checks to minimize biases of selection on observables. The dataset utilized comes from fieldwork conducted in early 2022 with key stakeholders from the Tanzanian cashew nuts sector. The quantitative primary data was collected with farmers following a multistage random sampling procedure. It was found that farmers connected to processing firms are associated with higher welfare outcomes when compared to those that sell only in the traditional marketing channel, which in this case, are government-run auctions. Specifically, when compared to SSFs that use exclusively the government-run auction system, those that engage with CF increased their output and reduced food insecurity by about 0.13 and 0.4 standard deviations, respectively. Additionally, processing firms face as binding constrains access to cheap credit, access to international buyers, and a steady flow of raw materials. The possible transmission channels identified are reductions in price and marketing risk, thus striking a dialogue with the latest applied CF literature. This chapter offers a compelling account of the development of the cashew nuts sector in Tanzania by identifying opportunities for value chain upgrading and specific challenges that need to be addressed.

Chapter 4 departs from a country-level focus to the global level, with a cross-country analysis. Agrifood markets have many inefficiencies that arise because of high transaction costs; increasing non-tariff measures (NTMs), poor transport and digital infrastructure, and burdensome trade procedures are some of the most important drivers of low productivity of the agrifood sector. Improvements in the digital infrastructure of customs and trade systems are increasingly key to cut trade transaction costs. Many countries have been facilitating international trade via the digitalization of bureaucratic processes and documentation. In recent years, a growing number of NTMs have been implemented to govern international agricultural trade. NTMs increase the bureaucratic burden on exporting firms and increase transaction costs, particularly in LMICs that have limited institutional capacity. In these settings, paperless digital trade offers the potential to facilitate compliance with NTMs as well as expedite other trade bureaucracies. Furthermore, many countries have initiated the facilitation of international trade bureaucracy via digital processes and documentation in line with the World Trade Organization (WTO) Trade Facilitation Agreement (TFA). In this sense, this chapter captured trade and customs digitalization with two main variables of interest: general paperless trade (PT) and digitalized NTMs compliance (CB). Cutting transaction costs is the mediation between facilitation via trade digitalization and higher agrifood trade because it streamlines trade and customs operational procedures. In order to test this hypothesis, I measured the effects of trade digitalization on bilateral agrifood trade using a representative cross-country dataset with a gravity model methodology, extending it with country- and time-fixed effects as well as socioeconomic controls, and estimated with the Pseudo-Poisson Maximum Likelihood (PPML) estimator. A one-standard deviation improvement in e-trade facilitation at the exporter level and in NTMs digitalization at the importer level increased agrifood exports by 9.7% and 1.3%, respectively, which is a high-bar finding. Specifically, I found that the positive trade effects of PT and CB are larger for processed products and for Sub-Saharan African and Asian countries. Consequently, trade facilitation via customs digitalization is inevitable to avoid trade diversion in an increasingly competitive globalized world. There are ample opportunities for LMICs agrifood exports to benefit from trade digitalization, because the relationship is stronger between agrifood exports and e-trade facilitation at the origin, which increases the agency of countries in the Global South. This could be especially achieved by adapting policies to the local realities, thus reducing implementation costs associated with technical challenges.

CHAPTER 2

Effects of agricultural market linkages on small-scale farmers' welfare: Evidence from Tanzania[17]

2.1 Introduction

Despite absorbing a large amount of labor, food crop agriculture in LDCs accounts for a proportionally low share of the GDP (Gollin et al., 2014a). In the specific case of Tanzania, it accounts for approximately two-thirds of the workforce and 26% of value added in the GDP, respectively[18]. Maize in LDCs is mostly cultivated on small plots of land tilled by hand and mainly rainfed, with low usage of marketable inputs and labor other than the family's own (Graeub et al., 2016; Lowder et al., 2016), with a similar outlook in Tanzania (Mmbando et al., 2015; NBS, 2021). Therefore, food crop agriculture is a mainstay of livelihoods in SSA, but key issues need to be addressed to improve rural development: poor infrastructure, which hinders access to markets (Stifel et al., 2016); meagre profit margins due to monopoly power from local traders (Kopp & Sexton, 2021); low utilization of modern inputs and technology, and low knowledge application into production techniques, which lowers productivity; limited access to capital, credit, energy, and irrigation; and weak governmental support (Collier & Dercon, 2014; Markelova et al., 2009). This study focuses on issues related to different forms of access to markets that can reduce transaction costs in the Tanzanian maize sector.

Insofar as 60% of the LDCs' poor are SSFs, while only 35% of the world's food is produced by them (Lowder et al., 2021), the poorest livelihoods in rural areas are inextricably linked to

[17] **Publication status:** Bueno Rezende de Castro, A., Kornher, L., Rugaimukamu, K. (2023): Under review for publication as a journal article.
[18] Figures are from the World Development Indicators from the World Bank.

the smallholder sector (Christiaensen et al., 2011; Rada & Fuglie, 2019) – poverty in the Tanzanian maize sector is widespread (NBS, 2021). Moreover, due to high population growth, SSA must increase yields far above the gains achieved in the last decades, in order to ensure food security, moving beyond extensive growth by adding factors of production (Diao et al., 2018). Otherwise, it will be a challenge for ordinary diets to feed people, especially in a context of diets shifting away from starchy staples towards the consumption of animal-based food products, fruits, and vegetables (Fukase & Martin, 2020).

The consequence of these problems is that the SSF sector, especially of food crops, faces high transaction costs that lead to worse welfare outcomes. To investigate if and which forms of agricultural market integration can reduce such transaction costs, I embed this chapter into two literatures: (1) how crop marketing improves SSFs' welfare, and (2) how agricultural market linkages are a key step from semi-subsistence towards commercialized production and, thus, the modernization of Tanzania's food system is a necessary condition to reduce rural poverty (Bellemare et al., 2022). I then provide evidence from Tanzania, showing that improvement of agricultural value chains through stronger linkages with marketing channels between SSFs and upstream actors is important to increase yields and income (thereby reducing poverty levels) as well as improving food security at the household level (Ogutu et al., 2020; Ogutu & Qaim, 2019). The focus of this chapter is twofold: (1) Tanzania is selected because of its strategic position as an East African trade hub, its uneven economic growth in the last decades with still high levels of poverty and undernourishment (FAO et al., 2020), and its similar structural characteristics to those of other SSA countries. (2) The crop focus is on maize due to it being the number one food crop consumed in the country and its potential to be cultivated in large scale targeting the agro-processing industry and export markets (NBS, 2021).

This chapter makes one major and one specific contribution to the literature. First, I provide empirical evidence for the effects of participation of maize farmers in different VMLs on rural

Tanzanian households' welfare (income, poverty, food security, and subjective well-being) as well as agricultural yields and commercialization. The methodology employs panel estimators to control for unobserved time-invariant covariates in conjunction with PSM to minimize sample selection bias from observed factors and to check robustness. Second, I provide a more comprehensive impact analysis of the effects of connection to marketing channels by using a richer set of welfare variables. Hitherto, most studies have analyzed one specific outcome variable, e.g., income, food security, or production indicators (Barrett et al., 2012; Bellemare, 2012; M. Maertens & Vande Velde, 2017; Mwambi et al., 2016; Nguyen et al., 2015; Soullier & Moustier, 2018), with the exception of a study on rice in Benin (M. Maertens & Vande Velde, 2017). The majority of studies focuses on separate production and linkage arrangements, i.e., CF, cooperatives, vertical integration, or only commercialization. This chapter addresses this gap in the literature with respect to a more comprehensive and structured analysis of such arrangements.

The remainder of this chapter is organized as follows. The next Section 2.2 provides a brief description of the main characteristics of the small-scale agricultural sector in Tanzania. Section 2.3 introduces the research hypotheses that are embedded in the context of a conceptual framework based on a literature review on SSFs and on the benefits and limitation of different forms of access to markets. Then, Section 2.4 describes the dataset employed and the measurement of the outcome variables, whilst Section 2.5 describes the measurement of the key variable of interest (VMLs) and presents the empirical strategy. Section 2.6 analyzes the main results, whilst Section 2.7 discusses these results in light of the applied literature. The last Section 2.8 concludes.

2.2 The small-scale agricultural sector in Tanzania

Sub-Saharan Africa continues to be the region in the world with the most pressing challenges in terms of food security; a situation that has only aggravated with the COVID-19 crisis and the ensuing spike of food and fuel prices, especially after the outbreak of the Russo-Ukrainian War. The sub-continent has the lowest scores in rankings of prevalence of undernourishment (undernutrition and subnutrition, SDG 2.1.1), access to food, per capita consumption of calories, prevalence of moderate or severe food insecurity (SDG 2.1.2), and anthropometric measures of children (SDGs 2.2.1 and 2.2.2) (FAO et al., 2022). These predicaments are explained by poor agricultural productivity and low incomes, which are further aggravated by the fact that a healthy diet costs over 85% of the average daily per capita household income in SSA and slightly more in Tanzania (Hirvonen et al., 2020). It is, thus, possible to conclude that food security is tightly interconnected with the development of SSFs and subsistence farmers in a LIC such as Tanzania.

Now turning the focus to subsistence farmers in Tanzania, they are particularly vulnerable to food insecurity because they: (1) often fetch paltry prices for their produce (after factoring in the relatively lower commercialization rate); (2) have low agricultural yields leading to meagre harvests; (3) are more often located in remote areas with higher transportation costs; and (4) have lower educational achievements and job experience, thus, fewer opportunities for off-farm employment (Fischer & Qaim, 2012a). Therefore, the issues these farmers are facing put the country at risk of failing to achieve the SDGs related to food security.

Among the main causes of these poor outcomes are the inefficiencies plaguing the small-scale agricultural sector in Tanzania, resulting in crops of poorer quality and lower value addition. Chiefly among these inefficiencies are market discoordination between producers, processors, and consumers, poor logistics (bad transport infrastructure and substandard storage facilities), low intensity in the usage of inputs, and bad practices in agronomy (Mutabazi et al., 2015; Putterman, 1995). The GoT implemented two main policy frameworks to address these issues, both named Agricultural Sector Development Policy: the first version was in effect from 2006 to 2014, and the second was enacted in 2017, effective until 2028. ASDP II differs from its predecessor in the sense that it has clearer targets and goals related to the enhancement of SSFs commercialization and income, with repercussions on improved food security and nutrition (Government of Tanzania, 2015). Notwithstanding this policy improvement, measurement, follow-up, and assessment of these targets and goals remain precarious.

2.3 Background and conceptual framework

First, I start this section with a brief discussion of the concept of transaction cost whereby the variable of interest (VML) is framed into the conceptual framework (see Figure 2.1).

Transaction costs and CF as a subcategory are two relevant aspects of VMLs. Firstly, selling crops via contracts is usually beneficial in environments featuring poor public institutions and market failures that disrupt the integration between supply and demand, taking a toll on market participants through higher transaction costs, making room for Pareto improvement (Grosh, 1994). Therefore, formal commercialization of crops, which is an integral aspect of VMLs, is a governance innovation that aims to reducing the still sizable transaction costs, particularly for SSFs (J. Swinnen & Vandeplas, 2012). Persistently high transaction costs induce lower net prices and marketable surpluses for SSFs, and higher food prices for consumers.

Secondly, the larger category of VMLs also includes CF, which are based on a contract between farmers and buyers. Similar to CF schemes, the legal and contractual environment in which farmers and agro-processing firms operate is important for VMLs proper operation (von Braun, 1995). Strictly speaking, contract farming schemes are a written (but often only verbal) rule regulating production and marketing relationships between farmers and processors. In Figure 2.1, VMLs overcome co-ordination failures caused by traditional spot markets – dominant for traditional and cereal crops (Feyaerts et al., 2020) –, playing an important role in modernizing the agricultural sector and improving welfare in rural areas.

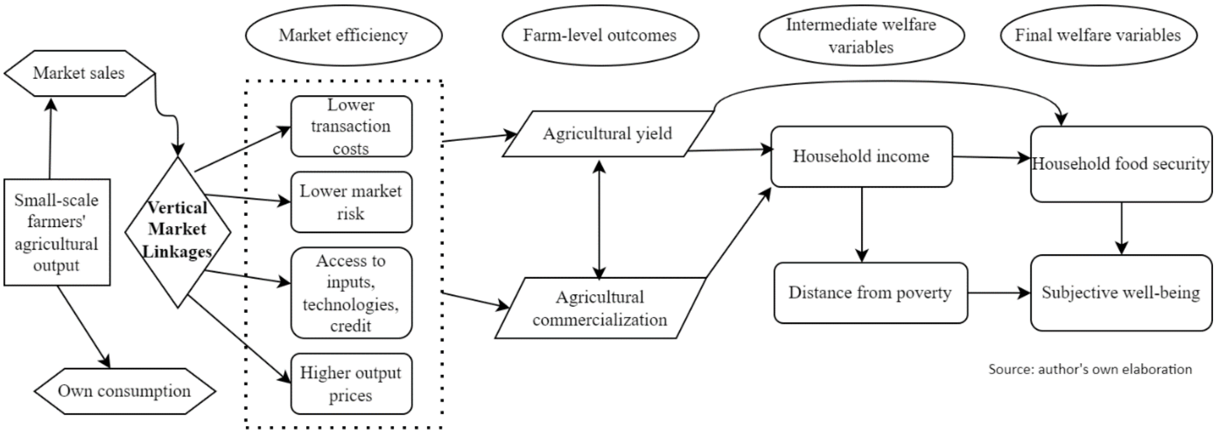


Figure 2.1: Conceptual framework of Chapter 2

The transition from subsistence to commercial agriculture is crucial within the modernization of food systems (Reardon & Timmer, 2012). The main factors thwarting the transition of SSFs from subsistence to commercial production are price uncertainty (Bellemare et al., 2013), access to credit (Berg, 2013), and a lack of technical knowledge. These constraints affect input demand, as well as yields, sales, and income, contributing to the perpetuation of backwardness in rural areas.

Agricultural commercialization explores specialization in certain crops and tasks along with comparative advantages, increasing agricultural productivity and total output (von Braun, 1995). Contract farming is quickly becoming a common marketing scheme, though less so in the maize sector (Ragasa et al., 2018), with pre-defined prices and quantities, but less so for production processes and input provision. Not only the potential of such schemes but also of VMLs as defined in this study resides in addressing important market failures: poor access to output markets, high transaction costs, high crop production risks, abundant information asymmetry, and poor access to inputs, credit and new technologies (Bellemare & Bloem, 2018; Otsuka et al., 2016). This often results in higher intensity of input use, better yields, and potentially higher incomes for SSFs (Bellemare & Novak, 2017). CF schemes are included in VML2 in the framework, as further explained in Section 2.5.1.

The conceptual framework follows the literature on food systems' transformation postulating that a higher degree of rural household integration into agricultural value chains can impact poverty reduction and shared prosperity (Reardon & Timmer, 2012; Tschirley et al., 2015). The main relationships between outcome variables and their driving factors are captured by the conceptual framework, whose core is the VMLs' effect on rural household welfare. Some of the market efficiency variables cannot be measured directly using data from the NPS-LSMS, apart from access to inputs, technologies, and credit, which are control variables, and output prices (not available for some VMLs). This pathway's theory of change explains how VMLs connect to outcomes. First, they minimize market failures, captured through market efficiency aspects. Agricultural markets function more efficiently with reduced transaction costs, lower market risks, and better access to inputs, technologies, and credit, resulting in higher output prices, as opposed to simple subsistence, thus benefiting SSFs (von Braun, 1995). Secondly, these efficiency gains directly boost agricultural yields and crop commercialization, thereby increasing rural household income and enabling SSFs to escape poverty. These comprise intermediate welfare variables – the indirect pathway. Last, households do not derive more utility strictly from higher income, but from the higher consumption it commands (Lancaster, 1966), thus improving final welfare variables, namely food security and subjective well-being – the direct pathway.

This conceptual framework builds on the literature attributing the benefits of CF to SSFs: (1) lower market risks by ensuring a market for the produce (Bellemare & Bloem, 2018); (2) easing of credit constraints when contracts provide loans and/or inputs (Casaburi & Willis, 2018; Otsuka et al., 2016); (3) improved access to higher quality inputs and technologies (Mishra et al., 2016); and (4) higher crop/output prices. Regarding the third benefit, a constant supply of high-quality agricultural raw materials is essential for developing an agro-industrial sector that is competitive in export markets (Chege et al., 2015). These are important for supermarkets catering to domestic demand, and the agro-processing industry, whose firms must fulfill many product grades, quality specifications, and technical requirements to qualify as international suppliers. There is still ample space both in the domestic and the international market for processed maize products other than flour; this can open opportunities for SSFs to connect to VMLs and overcome the overreliance on subsistence agriculture. Arrangements under the VMLs' rubric can deliver these standards and, therefore, bridge the gap between agricultural production and food processing. In this sense, the main question that this chapter addresses is whether SSFs become better off through such market linkages.

As aforementioned, many studies have concluded that CF schemes have positive welfare effects for SSFs. Other studies have found that improved market linkages, which enhance agricultural yield (Mishra et al., 2016) and improve income (Ton et al., 2018), might not necessarily lead to food security and better nutrition under certain conditions (Fischer & Qaim, 2012a). Therefore, we explore other outcome variables associated with household welfare, such as food security (Bellemare & Novak, 2017; Chege et al., 2015; Miyata et al., 2009), poverty (Ogutu et al., 2020; Ragasa et al., 2018), and subjective well-being (Dedehouanou et al., 2013; Väth et al., 2019), captured as final welfare variables. Some recent studies addressed the heterogeneous effects of different production schemes (Ochieng et al., 2017; Ruml & Qaim, 2020), however a gap remains regarding how different forms of marketing channels, as captured by VMLs in this study, impact SSFs welfare and production, especially in the Tanzanian case (Herrmann et al., 2018).

2.4 Data

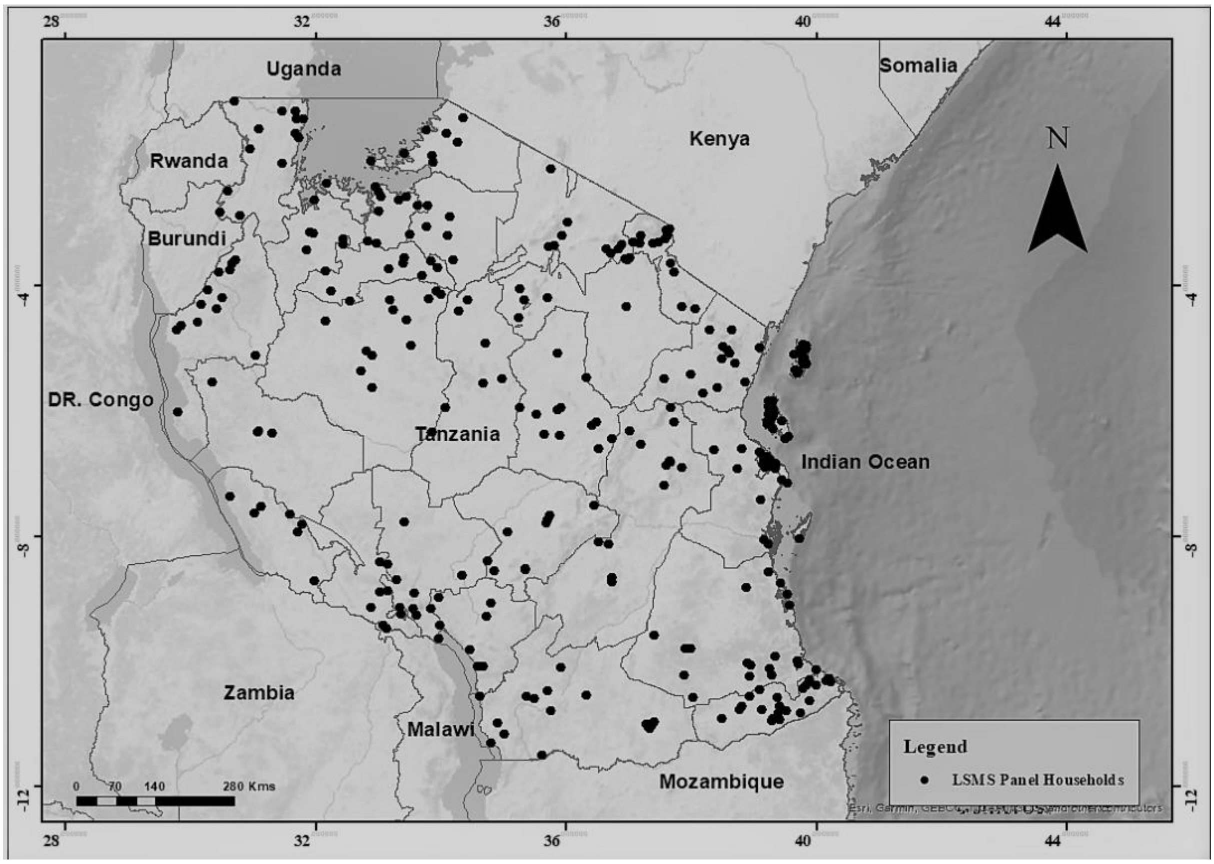


Figure 2.2: Locations of NPS-LSMS households in Tanzania

Source: Author's own elaboration based on NPS-LSMS data layered onto a Tanzania shape file using *QGIS*.

The secondary microdata source used in this chapter is the NPS-LSMS collected by Tanzania's National Bureau of Statistics and the World Bank's Living Standards Measurement Study - Integrated Surveys on Agriculture (LSMS-ISA). This dataset is representative of all rural areas and agro-ecological zones of the country; data on agricultural production was collected at the plot level, data on harvest at the household level, in addition to household and community characteristics. Contrarily to studies that have examined agricultural market linkages in the context of multiple crops (Miyata et al., 2009; Narayanan, 2014; Simmons et al., 2005), I focus on a single crop, which results in more accurate estimates. This chapter considers maize due to its key role in Tanzania as a food crop, produced by over 60% of households, and its expanding role as a cash crop (NBS, 2021).

I merge the first three waves of 2008–2009, 2010–2011, and 2012–2013, because the last two waves adopt a different sampling methodology, thus I prefer to rely on a balanced panel at the household level, rather than extending the panel in years but with different households. The sample size achieved is 1,502 households per wave, whose locations are marked in Figure 2.2 – the food security variables were not collected in the first wave of the NPS-LSMS. Considering the objectives of the study, I considered in the panel only farming households, i.e., those that owned or cultivated agricultural land in the years when data collection took place. There are only small systematic differences between farmers who leave the survey and those who remain, so the attrition bias is low. I tested for attrition bias by comparing the means of the outcome variables between the balances and unbalanced samples and they turned out to be similar. The few plot-level variables were matched with the respective households and transformed to household-level variables using unique household identifiers, given that the identification of the plot manager is available (Carletto et al., 2017). Table 2.1 summarizes the variables' definitions, omitting those fixed in the time span of the panel (e.g., gender, average distance to market/road, years of education, age), since they would be dropped by the panel estimators.

Table 2.1: Construction and definition of the variables

Dependent/Outcome Variables

Income, natural log	All household income sources (Tanzanian Shillings, Tsh)
Poverty	Dummy indicating households below the international poverty line
Yield, natural log	Main crop yield (kilograms/acre)
Commercialization	Total gross crop sales / total income
Food security	Household food insecurity access scale (HFIAS); household dietary diversity score (HDDS)
Subjective well-being (scale 0-7)	Score measuring households' general satisfaction and specific situations, i.e., financial, health, housing, job

Independent/Control Variables

Vertical market linkages	Two dummies for each of the crop sales' market channels
Wave	Three dummies for each NPS-LSMS wave, with wave 1 being the reference
Off-farm	Dummy: household head's off-farm employment
Plot use	Dummy: all plots being cultivated
Farm size (acres)	Farm plot area
Squared farm size (acres^2)	Squared farm plot area
Input credits	Dummy: household received inputs' credit
Organic fertilizer	Dummy: organic fertilizer use

Inorganic fertilizer	Dummy: inorganic fertilizer use
Intercropping	Dummy: intercropping practice, involving two or more crops
Pre-harvest losses	Dummy: farmers with pre-harvest crop losses
Post-harvest losses	Dummy: farmers with post-harvest crop losses
Labor hiring	Dummy: farmers who hired external labor
Health Shocks[19]	
Shock illness	Dummy: household member (HM) chronic/severe illness
Shock death	Dummy: HM death
Shock other death	Dummy: other family member death
Natural Shocks	
Shock drought/flood	Dummy: drought and/or floods
Shock disease/pest	Dummy: crop disease and/or crop pest
Shock water	Dummy: severe water shortage
Shock fire	Dummy: fire in household's dwelling
Economic Shocks	
Shock HH break	Dummy: household breakup
Shock land	Dummy: land loss
Shock crop price	Dummy: large fall in crops' sales prices
Shock input price	Dummy: large rise in agricultural input prices
Shock food price	Dummy: large rise in food prices
Shock livestock	Dummy: livestock death/theft
Shock business	Dummy: non-agricultural business failure
Shock salary	Dummy: loss of salaried employment and/or non-payment
Crime/Conflict Shocks	
Shock jailed	Dummy: HM jailing
Shock robbery	Dummy: household dwelling's hijacking/burglary/robbery/assault
Shock dwelling	Dummy: household dwelling's destruction/severe damage
Shock other	Dummy: any other shock types

2.4.1 Outcome variables

The measurement and description of the outcome variables are defined as follows:

1. Household income: measured by all income sources (agricultural and non-agricultural) captured by the NPS;

2. Income poverty: measured by a dummy variable for households below and above the international poverty line[20], using household income as reference (Ogutu et al., 2020);

3. Agricultural productivity: measured by the plot-level yield per acre of maize proxying the household yield;

[19] Definition of shock groups adapted from Brück et al. (2019) and Smith and Frankenberger (2018).
[20] The World Bank's international poverty line (US$ 1.90 per person/day in 2011 PPPs) is more restrictive than the national one.

4. Crop commercialization: measured by the ratio of the gross value of sales of all crops produced to total income. The commercialization variable captures the degree to which a given household's income is connected to agricultural markets (von Braun (1994)). The variable for crop commercialization transcends a simple dummy variable that only captures the dichotomy between being strictly a seller and strictly a subsistence farmer (Carletto et al., 2017);

5. Subjective well-being of household head: measured by the arithmetic mean, based on a seven-point Likert scale of responses about the household satisfaction level with various life components: personal health, healthcare availability, financial situation, employment, housing, education availability, safety, and life as a whole[21].

6. Food security: measured by the standard indicators HFIAS and HDDS as defined by Coates et al. (2007). The HFIAS index records the household's self-reported food consumption over the last seven days, including food produced, purchased, or received as gifts/donations, defined by the purchase value or the average crop price that the household received if self-produced. The HDDS is an intrinsically qualitative index transformed into a quantitative score composed of the number of different food groups consumed over the last seven days, with one point for each additional food group consumed by the household. It captures more nuances (i.e., nutritional quality) than HFIAS and reflects a households' economic ability to access different types of food items, thus an increase in HDDS is associated with higher welfare (Abdulai & Aubert, 2004; Kennedy et al., 2010). The two indices combined offer a good characterization of the canonical food security definition (CFS, 2012): (1) availability of sufficient food amounts; (2) sufficient purchasing power for food consumption; (3) sufficient quality food for a nutritious diet; and (4) stability of food availability. Food security is a particularly important variable; opportunities for paid work in rural areas being scarce, many resort to subsistence agriculture (Barrett, 2010). One limitation of these two indexes is that they do not capture intra-

[21] The exact question for each of these components reads as follows: "How satisfied or dissatisfied would you say you are with…" and the respective category. The responses vary from "very satisfied" to "very dissatisfied" and are assigned a corresponding number from 1-7.

household food distribution.

2.5 Methodology

This chapter's methodology adopts three stages:

1. I start with the ordinary least squares (OLS) estimator to provide a benchmark;
2. I conduct an analysis using panel data to identify the impact of VML participation on outcome variables, addressing potential endogeneities, especially from unobservable covariates (see Tables 2A.3-2A.5). Although the panel estimators enable a better isolation of the VMLs' effects on the outcome variables and despite the several models and estimations to test for robustness, I remain cautious about causality and prefer to talk about associations.
3. As a robustness check, I employ PSM separately for each wave and VML group in order to limit observations between comparable groups, since I drop the observations with unbalanced matching (see Table 2A.6). The estimated average treatment effects on the treated (ATT) are then less prone to self-selection on observed covariates, resulting in a reduction of selection bias (see Tables 2A.7 and 2A.8).

2.5.1 Vertical market linkages: measurement and description

I employ Herrmann et al. (2018)'s methodology to represent different market participation channels, with VML being the variable of interest. Essentially, VMLs capture different market channels for agricultural products, excluding produce sold on spot markets, to neighbors, to friends/family, or consumed in the household (e.g., subsistence) – the control group. Hence, I create a binary classification of market participation channels with two groups based on farmers' sales[22]:

[22] Definitions based on the market channels from the NPS-LSMS' network roster card.

1. VML1 group: Cooperatives and farmers' organizations/associations + Grocery/Local merchant + Private business people. Here I group the traditional cooperative market channel, through which farmers' groups can leverage market power and improve profit margins, with less traditional market channels. The last two market channels are not involved with AP firms, thus agricultural produce is purchased with informal arrangements, often via traders.
2. VML2 group: Business contacts + Employer + Contract Farming + Factories and slaughterhouses. This group represents direct connections with agro-processing firms via formal channels – CF, which has a small sample size in the NPS-LSMS, and factories – and quasi-formal channels, which are sales to local businesses with less formalized contracts.

I employ the VML variable to represent the agricultural development in SSA via different vertical integration forms with supply chains (Barrett et al., 2022; Bellemare et al., 2022; Kirsten & Sartorius, 2002). This category focuses on produce sold to AP firms or individuals who process raw materials and sell the goods with higher value added. The objective is to capture the processing of raw materials and its value chains' integration, which is the key characteristic that distinguishes VML from subsistence farming and selling to local spot markets. Lastly, the farmers that are not classified neither on VML1 nor on VML2 are considered as not having market linkages, and, therefore, they serve as a subsistence benchmark of the outcome variables.

For the purposes of this dissertation, the categories of VMLs capture more sophisticated forms of market organization, compared to subsistence farming, that cut transaction costs for rural producers. This reduction is associated with higher market efficiency and ultimately greater household welfare. In this sense, this chapter does not precisely estimate the size of transaction costs in each type and sub-component of VML but rather estimates the effects of VML participation on welfare outcomes. These, in turn, serve as indirect measures of how lower transaction costs can improve rural households' welfare.

2.5.2 *Fixed effects vs. correlated random effects*

Here I present the panel regression models employed and discuss how they correct for potential endogeneity sources in the estimators. The main model estimated using fixed-effects (FE) generalized least squares is as follows:

$$Y_{it} = \beta_{FE,0} + \beta_{FE,1}VML1_{it} + \beta_{FE,2}VML2_{it} + \beta_{FE,3}(VML1_{it} \times VML2_{it}) + \beta_{FE,4}X_{it} + \gamma_i + \delta_t + \varepsilon_{it},$$

(2.1)

where the subscript *i* indexes farming households and *t* indexes time (the NPS-LSMS wave). I estimate six separate models for each dependent variable to avoid simultaneity bias; however, for simplicity, only the main structural model is presented. *Y* represents six separate outcome variables (see Table 2.1). I also add the interaction term β_3 for the two VML channels absorbing bias from the few observations common to both groups (i.e., households selling their produce in both market channels). Correspondingly, the Wald F test rejects the H_0 that the variables of interest are jointly not significant.

X is a vector of time-variant control variables expected to affect the outcome variables (see Table 2.1). Assumed as time-invariant, household characteristics (i.e., household size, household head's years of education, age, and gender) are dropped, being cancelled out by means differencing. The time-invariant unobservable household characteristics are captured by γ, and the panel/time effects are represented by δ, capturing structural changes in the time span of the panel, such as economic growth, infrastructure improvements, and other shocks not included in vector X. I cluster the standard errors ε_{it} at the household level for three reasons: (1) to make them robust to heteroskedasticity and serial correlation (Wooldridge, 2010); (2) the assignment to VML could be correlated with clusters (Abadie et al., 2022); and (3) to correct for any loss of degrees of freedom when estimating N individual means (Angrist & Pischke, 2009).

The advantage of the FE over the random-effects (RE) model is that the former does not assume that the unobserved time-invariant component γ is not correlated with the control variables: Cov (β_i, X_i) ≠ 0; the FE estimator corrects this correlation when subtracting the values for each wave from the group means, thus cancelling out γ (Wooldridge, 2010). In addition to the FE model, I also adopt the Mundlak approach (or correlated random effects, CRE) which has two main advantages over regular panel data estimators: first, it produces consistent estimates compared to the RE model because it does not assume that time-invariant unobserved variables are uncorrelated with the control variables; second, it produces more efficient estimates than FE if the within variation in the data is smaller than the between variation, which is our case (see Table 2A.2); third, it includes the means of the time-varying controls and VMLs with the goal to reduce selection bias stemming from unobserved household heterogeneity (Wooldridge, 2019). In this sense, both the CRE and the FE estimator correct the bias from assuming that farmers self-select into VMLs on unobserved but fixed individual characteristics, mainly motivation and farming ability. The choice between these two estimators is determined by the Mundlak test: if there is correlation between time-varying variables and the time-invariant component of the error term, I prefer the CRE specification; if not, then we prefer the FE specification. The CRE model is defined as follows:

$$Y_{it} = \beta_{CRE,0} + \beta_{CRE,1}VML1_{it} + \beta_{CRE,2}VML2_{it} + \beta_{CRE,3}(VML1_{it} \times VML2_{it}) + \beta_{CRE,4}\overline{VML1}_{it} + \beta_{CRE,5}\overline{VML2}_{it} + \beta_{CRE,6}X_{it} + \beta_{CRE,7}\bar{X}_{it} + \gamma_i + \delta_t + \varepsilon'_{it}. \quad (2.2)$$

All variables are as defined in Equation 2.1; the main difference to the FE model is the presence of the coefficients β_4, β_5, and β_7, which denote the averages of the time-varying covariates. The coefficients of interest are β_1 and β_2, capturing the impact of participation in each of the VMLs on the measured outcome variables. Both are expected to be positive, except for the outcome variable measuring poverty and HFIAS, since VMLs have a welfare-increasing effect. The hypothesis to be tested is:

- H_0: $\beta_1, \beta_2 \neq 0$ ➔ Parameters should have a magnitude greater than zero, with statistical significance.

For this estimate, I use data from a balanced panel controlling for time and geographical fixed effects. The panel methodology enables us to draw inferences with less bias than OLS because it is possible to control for entity heterogeneity. The removal of time-invariant household characteristics is then possible, thus controlling for time-fixed effects (within the panel time span) and noise from confounding factors (i.e., independent variables and other household specific characteristics). However, omitted variable bias could still arise from unobserved time-varying heterogeneity. Despite assuming that this factor is minimal compared to the time-invariant household characteristics, I still control for a wide array of socioeconomic shocks: off-farm employment, plot usage, credit received for inputs, agronomic variables, and farm size. Another potential simultaneity bias between outcomes and the variable of interest is an endogeneity source that can never be completely ruled out without suitable instruments: the decision of where to market the produce (if at all). In the conceptual framework, I put forward a rationale that explains why it influences SSF's welfare.

2.5.3 First differences

The advantages of FD estimators are: first, cancelling out unobserved time-invariant characteristics, but in a different way regarding FEGLS; by keeping all households, not only those who changed participation status, assimilating the effect of households that always participated in VMLs and those that never did. Secondly, the FD approach relaxes the assumption of a non-dynamic relationship between treatment and outcome variables, with different treatment and control groups (i.e., those who always/never had VML and those who changed).

The model estimated using FD is defined as follows:

$$\Delta Y_{it} = \alpha_1 \Delta VML1_{it} + \alpha_2 \Delta VML2_{it} + \alpha_3 \Delta X_{it} + \Delta u_{it}. \tag{2.3}$$

Variables with Δ are measured in FD, subtracting the previous period from the last, rather than in levels, losing one observation. The coefficients of interest are α_1 and $\alpha_{2,}$ measuring the effects of VMLs in FD on the outcome variables. X is the same vector of time-variant control variables.

The main intuition is that there are more changes in outcomes for households participating in VMLs than for those who do not – mainly subsistence farmers. This is a direct implication of the parallel/common trends assumption: without participation in VMLs, the average change in outcome variables would follow the same trends for both SSF groups – participants and non-participants. The change could still lead to disparate outcomes, but the trends would roughly be equal. The two reasons a common trend should be assumed for all households are: (1) lack of major shocks that severely affected the Tanzanian agricultural GDP in the NPS-LSMS waves' period[23], and (2) lack of major shocks not covered by the controls that could have disproportionately impacted a specific subsample.

2.5.4 Propensity score matching

Lastly, the PSM analysis aims to control for an endogeneity problem that is particularly common in studies using observational data: selection bias (Barrett et al., 2012). As VML participation is likely not to be random, selection bias occurs with both observable and unobservable variables: farmers who have better socioeconomic conditions, lower risk aversion, higher innate skills, and more motivation for farming tend to generate higher yields (Barrett et al., 2012; Chege et al., 2015). When selecting suppliers, whether contractually or not, buyers often exclude those with revealed lower productivity, fewer skills, less access to

[23] Data from the World Development Indicators (2008-2020).

resources, and farther from plants (M. Maertens & Swinnen, 2009; Rao & Qaim, 2011).

First, I match the treatment group (VMLs) with the control group on propensity scores estimated based on a vector of observable characteristics, thus reducing selection bias in the panel models' sample (Gibson & McKenzie, 2014). Then I estimate the determinants of VML participation using a binary Probit model (Imbens, 2004); last, the robustness check uses three matching techniques with replacement (i.e., nearest-neighbor, kernel, and radius) to estimate the ATTs, which are the difference between the mean outcome of participants and the mean outcome of similar non-participants (Caliendo & Kopeinig, 2008). As sensitivity checks, I employ the bounding approach to estimate the influence of unobserved factors relative to observed factors, ascertaining the PSM bias to unobservables (Ichino et al., 2008; Rosenbaum & Rubin, 1983). Additionally, I use relevant socioeconomic and agronomic variables to control for as much characteristics as possible driving the estimates (Heckman et al., 1997). Nevertheless, a clear limitation of PSM is that part of the differences in outcomes might result from non-observable factors predicting VML participation, rather than the effect of VMLs or observable controls (Soullier & Moustier, 2018).

2.6 Results

2.6.1 Descriptive statistics

This sub-section starts by comparing means using a t-test for households participating in vertical market linkages versus non-participants, separating them into two VML groups (see Tables 2.2 and 2.3). Significant differences are identified between participants and non-participants in all dependent variables – except for subjective well-being -, showing that the two samples do indeed differ. Nevertheless, I observe lower statistical significance for VML1 because of the smaller sample size.

Table 2.2: Descriptive statistics by NPS-LSMS wave for VML1

	Mean (SE)		Diff	Mean (SE)		Diff	Mean (SE)		Diff
	W1(0)	W1(1)	1 - 0	W2(0)	W2(1)	1 - 0	W3(0)	W3(1)	1 – 0
Outcome variables									
Income (log)	12.92	13.97	***	13.22	13.98	***	13.62	14.32	***
	(0.046)	(0.149)		(0.046)	(0.114)		(0.043)	(0.107)	
Poverty (dummy)	0.875	0.848		0.851	0.728	***	0.777	0.675	***
	(0.009)	(0.054)		(0.10)	(0.035)		(0.011)	(0.038)	
Commercialization (0-1)	0.772	0.866	*	0.737	0.804	**	0.687	0.795	***
	(0.010)	(0.040)		(0.011)	(0.026)		(0.011)	(0.023)	
Crop yield (log)	5.12	5.31		4.92	5.24	***	4.98	5.05	
	(0.039)	(0.144)		(0.035)	(0.105)		(0.037)	(0.102)	
HDDS (1-10)	NA	NA		7.39	7.75	**	7.21	7.36	
				(0.049)	(0.139)		(0.053)	(0.168)	
HFIAS (score)	NA	NA		8.15	5.36	**	8.48	5.88	*
				(0.415)	(1.000)		(0.483)	(1.15)	
Subjective well-being (1-7)	3.84	3.92		4.00	3.98		4.04	4.05	
	(0.039)	(0.198)		(0.036)	(0.123)		(0.037)	(0.109)	
Observations	1,456	46		1,340	162		1,348	154	

Note: W = NPS-LSMS wave. Households are classified as either non-participants (0) or participants (1) in VML1. Mean values are shown with standard errors in parentheses. p-values: * p<0.10, ** p<0.05, *** p<0.01.

Table 2.3: Descriptive statistics by NPS-LSMS wave for VML2

	Mean (SE)		Diff	Mean (SE)		Diff	Mean (SE)		Diff
	W1(0)	W1(1)	1 - 0	W2(0)	W2(1)	1 - 0	W3(0)	W3(1)	1 – 0
Outcome variables									
Income (log)	12.92	13.80	***	12.96	13.71	***	13.29	14.13	***
	(0.046)	(0.186)		(0.065)	(0.053)		(0.060)	(0.050)	
Poverty (dummy)	0.880	0.714	***	0.861	0.808	***	0.835	0.690	***
	(0.009)	(0.061)		(0.012)	(0.015)		(0.013)	(0.017)	
Commercialization (0-1)	0.716	0.778		0.695	0.804	***	0.628	0.773	***
	(0.10)	(0.053)		(0.015)	(0.013)		(0.015)	(0.012)	
Crop yield (log)	5.04	5.13		4.98	5.43	***	4.91	5.16	***
	(0.039)	(0.236)		(0.047)	(0.045)		(0.052)	(0.047)	
HDDS (1-10)	NA	NA		7.31	7.58	***	7.07	7.40	***
				(0.064)	(0.066)		(0.073)	(0.069)	
HFIAS (score)	NA	NA		9.32	6.01	***	8.94	7.38	*
				(0.561)	(0.503)		(0.665)	(0.592)	
Subjective well-being (1-7)	3.84	3.80		4.05	3.94		3.97	4.11	**
	(0.039)	(0.207)		(0.047)	(0.052)		(0.048)	(0.051)	
Observations	1,446	56		834	668		796	706	

Note: W = NPS-LSMS wave. Households are classified as either non-participants (0) and participants (1) in VML2. Mean values are shown with standard errors in parentheses. p-values: * p<0.10, ** p<0.05, *** p<0.01.

Table 2A.1 shows the importance of controlling for shocks because the proportion of households that suffered each type of shock is high; most suffered at least one, with economic and natural shocks being prevalent. I also test the implicit hypotheses from the conceptual framework whereby graphs are plotted for the relationships between the dependent variables, all in the expected directions, except for subjective well-being (see Figures 2A.1-2A.7).

Table 2A.2 reports the unconditional probabilities of SSFs transitioning from one VML group to the other or back to subsistence (no VML). The transition matrix shows sufficient variation within VML groups over time, although this is larger for the processors group compared to the cooperatives group, including a higher probability of SSFs leaving the cooperatives' VML compared to the processors' VML, hence resilience in the latter is greater than in the former.

2.6.2 Panel regression results

First, I briefly report on the main covariates. Off-farm labor is negatively associated with commercialization, which suggests that farmers who have other work activities have less time to dedicate to farming and, thus, tend to sell less crops in markets; however, off-farm labor is positively correlated with income and negatively with poverty. Smaller farms tend to marginally have higher yields than larger ones and intercropping positively impacts yield and income.

Second, I apply PSM to reduce the sample selection bias. Eleven controls are excluded from matching until the balancing property is satisfied, which means that the controls included are the correct ones that match the treated and non-treated observations based on observed covariates. Thus, about a fifth of the sample is excluded ensuring that the remaining observations are appropriate for statistical inferencing. I conduct Hausman tests in which all outcome variables' coefficients are consistently estimated with the FE model vis-à-vis RE.

The coefficients summarized in Table 2.4 should be interpreted as the average marginal effect (AME) of participation in one of the market channels relative to the control group, which are subsistence SSFs (i.e., not connected to any VML) – each coefficient comes from separate regressions for each outcome variable including the same controls. Overall, the interaction terms between the two treatment variables (VMLs) are not significant, possibly because households generally sell their crops in one of the two VMLs, so I interpret the AMEs directly from the coefficients of the separate VMLs. There is some divergence in the effects of the two market channel groups, as the effect of SSFs taking part in VMLs is large and significant mostly for the processors group and all outcome variables, except for subjective well-being. The controls that reported significant impacts on the outcome variables are off-farm income, farm size, intercropping, fertilizer use, and natural shocks related to agricultural production. However, there is a caveat concerning the coefficients of the other household characteristics: many are not significant most likely because of insufficient within-household variations over time.

VML1's effect on commercialization is only significant in the OLS benchmark model and at the 10% level in the CRE model, whereas the effect of VML2 is significant in all models. In the FE specification, participation in VML2 increases commercialization by 8.67%[24] compared to the mean of non-participant farmers; the result is a bit higher in the CRE specification, whereas the percentage change for VML1 is similar, though lower. As expected from the conceptual framework, higher commercialization is associated with higher crop yields (see Figure 2A.1), and I observe a strong and significant result of participation in VML2 on yields: a 33% increase[25] using FE and a slightly higher value using CRE. For VML1 the effect is much smaller and is not statistically significant, apart from the OLS benchmark (see Table 2A.3).

Table 2.4: Summary of the effects of participation in VMLs on the outcome variables

Outcome variable	**VML1**				**VML2**			
Model specification	**OLS**	**FD**	**FE**	**CRE**	**OLS**	**FD**	**FE**	**CRE**
Log yield	0.262*	0.019	0.0132	0.113	0.661***	0.377***	0.283***	0.300***
	(0.141)	(0.119)	(0.163)	(0.157)	(0.086)	(0.073)	(0.081)	(0.080)
Commercialization	0.124***	0.038	0.042	0.052*	0.121***	0.059***	0.057***	0.060***
	(0.024)	(0.023)	(0.031)	(0.029)	(0.015)	(0.014)	(0.016)	(0.016)
Log income	1.093***	0.483***	0.529***	0.718***	0.887***	0.540***	0.530***	0.567***
	(0.141)	(0.161)	(0.204)	(0.199)	(0.085)	(0.098)	(0.112)	(0.110)
Poverty	-0.061***	-0.032	-0.044	-0.0488	-0.073***	-0.040***	-0.067***	-0.067***
	(0.029)	(0.030)	(0.036)	(0.034)	(0.015)	(0.015)	(0.017)	(0.017)
HDDS	0.290*	0.253	0.292	0.345*	0.261***	0.247***	0.256***	0.272***
	(0.155)	(0.167)	(0.210)	(0.193)	(0.073)	(0.078)	(0.083)	(0.081)
HFIAS	-3.547***	-1.289	-2.418	-2.852**	-2.255***	-2.169***	-2.437***	-2.714***

[24] To compute the percentage change, we consider how much the estimated coefficient represents from the sample mean.
[25] For computing the percentage change in the log-level models (i.e., income and yield), we use the Halvorsen-Palmquist correction.

	(1.064)	(1.334)	(1.571)	(1.435)	(0.633)	(0.879)	(0.959)	(0.894)
Subjective well-being	-0.020	0.181*	0.141	0.095	-0.035	0.101*	0.043	0.013
	(0.100)	(0.104)	(0.130)	(0.124)	(0.051)	(0.056)	(0.061)	(0.059)
Observations	4,506	3,004	4,506	4,506	4,506	3,004	4,506	4,506

Note: OLS: ordinary least squares; FE: fixed effects; FD: first differences. Significance levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Observations for HDDS and HFIAS: 3,004. Source: Author's own calculations based on the NPS-LSMS.

Next, turning to income, although about two-thirds of the households in the sample are employed in off-farm activities, most of their income still comes from agricultural production and crop sales, which indicates that the majority engage in informal or low value-added activities. The effect of participation in VMLs are the same for both groups in the FE specification (about 70% higher than the subsistence benchmark) and similar in the FD, whereas the effect is larger for VML1 considering the CRE estimator – it doubles (105% vs 76%). Both effects are reduced when moving from the OLS to panel models, again indicating attenuation of bias. Income is directly related to poverty as households commanding a higher income are farther from poverty (see Figure 2A.4). Nevertheless, I observe smaller effects, as participation in VML2 reduces poverty by 7.7% in both the FE and CRE models. The reduction in poverty from participation in VML1 vis-à-vis non-participant farmers is smaller and only statistically significant in the OLS benchmark (see Table 2A.4).

Regarding food security, the HFIAS's sign is negative because the conceptual framework expects that market channels' participation reduces food insecurity; analogously, I expect a positive sign for HDDS as participation in market channels is hypothesized to increase dietary diversity. VMLs' effects on HFIAS are greater than on HDDS, indicating that households with VML participation are better-off in terms of total food consumption (i.e., lower food insecurity), but not as much in terms of variety of food types consumed (see Table 2A.5) – the maximum effect is around an extra 0.3 food groups. Specifically, VML2 participation increases dietary diversity by 3.6% and 3.8% in the FE and CRE models, respectively, whereas VML1 participation has an impact on HDDS with lower statistical significance only in the OLS benchmark and in the CRE model (4.8%). Regarding the HFIAS index, compared to non-participants, VML2 participation reduces food insecurity by 25.5% and 28.4% in the FE and CRE models, respectively, whereas for VML1, the impact is only statistically significant again in the OLS benchmark and CRE model (29.9%). Lastly, VMLs participation only impacts subjective well-being with statistical significance at the 10% level in the FD model; not higher than 4.5% (see Table 2A.5).

Here I shed light on the results based on the conceptual framework, as shown in Figure 2.1. The VML groups affect the final welfare variables through their effect on farm-level outcomes and intermediate welfare variables. Although I observe that better integration with market channels, which results in higher commercialization rates, directly improves crop production, such improvements have higher statistical significance and magnitudes for VML2. The income derived from the household – both in agricultural production and in other activities – increases by a much greater magnitude for both VML groups, despite poverty showing a statistically significant reduction only in VML2. Food insecurity and dietary diversity also have higher statistical significance in VML2; it is also not surprising that the magnitude of the effect on HFIAS is much larger than on HDDS because households tend to first increase their purchase of staples to fulfill caloric needs before they increase their dietary diversity (Ameye et al., 2021; Van den Broeck et al., 2021). Finally, for the other final welfare outcome variable, subjective well-being, the effects are only significant in the FD model.

The interpretation of these results begin with two conclusions: (1) the heterogeneity of the two different market channels is key to understand the results; (2) the effects of participation on market linkages on the outcome variables are lower when the estimated model moves from OLS to the panel ones, indicating that the correction of the model specification reduces bias. Then, first, the positive impact of more sophisticated market access is captured by VML2's effects, which are greater and more statistically significant across the board when compared to VML1. Recalling Section 2.5.1, VML2 is composed of more formalized connections between farmers and AP firms compared to VML1, that tend to have a more resource-providing nature making it a potential transmission channel. Second, I acknowledge the limitations of VML1's smaller sample size, with the effects likely being driven by the cooperatives' channel, which tend to have underdeveloped market connections in Tanzania for food crops, such as maize. The lower variation of the VML1 dummy leads to higher standard errors in the FE estimates, which are partially corrected by the CRE estimator. Lastly, in places with low contract enforcement – a defining characteristic of LDCs –, SSFs tend to prefer traditional market schemes without a formal contract, thus placing a higher value on personal relationships based on trust (Blandon et al., 2010; Gelaw et al., 2016; Schipmann & Qaim, 2011),[26] despite frequently disappointing outcomes compared to the potential of formal markets. This might explain VML2's stronger results as well as its higher sample size relative to VML1.

The results are similar to those found by Herrmann et al. (2018) regarding the positive effect of vertical market integration on crop commercialization and incomes in Tanzania, however, regarding food security, the results turn significant. This might arise because I did not restrict the VML variable to only CF and cooperatives, categories that are poorly captured in the secondary data and with a small sample size, as reflected in the weaker effects of VML1. I found a clear effect of the processors' VML group on reducing food insecurity and increasing dietary diversity, however these effects have lower statistical significance for the cooperatives' VML group since only the CRE estimator captures such effect. In contrast to this aforementioned study, which focused on cooperatives/farmers' groups and CF, I found stronger and more significant effects in the processors VML group across the board for all indicators. Analogously, Mmbando et al. (2015) used smallholder's consumption as the outcome variable and found that participation in maize and pigeonpea markets in Tanzania also have a welfare increasing effect. Lastly, Ogutu & Qaim (2019) used crop commercialization as the variable of

[26] This is true unless the contract contains explicitly favorable conditions and clauses, which does not always occur (Ola & Menapace, 2020b).

interest, instead of outcome, and found an effect of market integration on poverty reduction with a larger magnitude than this chapter found.

2.6.3 Robustness checks

2.6.3.1 Panel models

Stepwise regressions[27] are also employed to test the sensitivity of the results to introduction of relevant control variables, for each dependent variable. First, the restricted model with VMLs only; second, the restricted model adding wave dummies; and third, the full unrestricted model adding controls in a stepwise manner. Results are confirmed across these different specifications. Lastly, given variance inflation factor values smaller than 10, I found that multicollinearity within the covariates is not an issue[28].

2.6.3.2 Propensity score matching

There are four main criteria to analyze the balancing tests (see Table 2A.6) that are as follows:

1. The LR chi-squared test should have a p-value higher than 0.1 so that I can fail to reject the null hypothesis that the model of variables explaining the difference in assignments is no longer significant as a whole;
2. The mean and the median bias of the matching should be lower than 5%;
3. The t-tests' p-values for the matching variables should be higher than 0.05. This indicates that there is no difference between the treated and untreated groups after matching;
4. The absolute standardized difference in the means and the variance ratio – Rubin's B and R, respectively – should not exceed the limits set by Rosenbaum and Rubin (1985).

[27] Full tables are available upon request.
[28] Results are available upon request.

The matching shows presence of common support using all three waves with the key VML variables (see Figures 2A.8-2A.9). Only very few observations are outside the common support region and had to be excluded from the matching, hence showing that selection bias was partially corrected. Then, I compute the ATTs of VML participation (i.e., treatment) on the outcome variables using the aforementioned matching techniques (see Tables 2A.7 and 2A.8). Except for subjective well-being, all ATTs are significant for VML2, and for VML1 no statistical significance is observed only in the ATTs for poverty, subjective well-being, and HDDS. And lastly, I followed Rosenbaum (2005) to check if the ATTs are biased due to the presence of unobservable covariates: the estimates that are statistically significant are robust to sufficiently high hidden biases. The Rosenbaum sensitivity test allows the validity of this assumption to be tested and is presented in Tables 2A.7 and 2A.8 – with the preferred matching algorithm being the Kernel one. Except for subjective well-being, I observe a high level of robustness to unobserved confounders in the processors' VML, however this robustness is quantitively smaller in the case of the cooperatives' VML. The values in the columns "critical hidden bias" show a factor to which each of the outcome variables are sensitive to the introduction of unobserved confounders. In general, the estimates are found to be robust or insensitive to an unobserved bias that would be at least 20% higher than the odds of non-adoption of VMLs. In short, the PSM methodology largely corroborates the validity of the panel estimations.

2.7 Discussion

Ragasa et al. (2018) found that, although CF arrangements for growing maize in Ghana resulted in higher yields and technology adoption, it did not increase farmers' profitability simply due to the transmission channel of higher input and credit costs. This provides evidence that sometimes CF schemes are not able to correct market failures, but rather need correction before the benefits from CF bear fruit. Naturally, this is beyond this chapter's scope, however their conclusion strikes a point to the finding that, although VMLs lead to higher incomes and enhanced production outcomes, they do not translate as clearly into higher household welfare in terms of food security and subjective well-being.

A limitation of this chapter could be considering only one quantitative measure of income poverty based on the international poverty line, which could affect the result that VMLs only reduce poverty in the processors' VML group. There is a large body of literature following Sen (1981) analyzing poverty from a more holistic angle: poverty cannot be measured only by disposable income, but the possibilities and restraints to access of resources ought also to be considered. This leads to an ample avenue for further nuanced studies on poverty and agricultural value chains, such as Ogutu & Qaim (2019), who collected primary data that allows for poverty to be measured in a multidimensional way.

Lastly, this chapter has three drawbacks regarding measuring food security since the dataset utilized did not allow for these: first, I do not analyze food security variables for the first wave; second, I do not consider micronutrients consumption, i.e., hidden hunger; third, I do not look at food security and food consumption at the individual level, only at the household level. Regarding the importance of analyzing nutritional outcomes, the theoretical framework elaborated by von Braun and Kennedy (1994) emphasizes the complexity of the interconnections between the external environment and SSFs, leading to nutritional outcomes. Although a higher income can improve nutrition, its effects are mediated by intra-household allocation of food and non-food expenditures following strict gender lines. Thus, I acknowledge the literature looking at food security at the individual level as well as the importance of micronutrients consumption for individual well-being and avoidance of "hidden hunger" (Abdulai & Aubert, 2004; Stein & Qaim, 2007), however the NPS-LSMS does not have data on either aspect. This is a good avenue for further studies.

2.8 Conclusions and policy implications

Sub-Saharan African farmers are increasingly selling more crops while subsistence consumption is declining in relative terms (Bellemare et al., 2022; Liverpool-Tasie et al., 2021), and these trends were not disrupted by the COVID-19 pandemic (Engemann & Jafari, 2022). How does this benefit small-scale farmers? This study examines the effects of small-scale farmers' integration to vertical market linkages on farm-level outcomes (crop yield and commercialization), intermediate welfare (household income and poverty), and final welfare (food security and subjective well-being), thus contributing to the literature on agricultural value chains. The findings are in line with the literature identifying a positive impact of more sophisticated market linkages on SSFs. However, the precise channels and the heterogeneity of effects according to VML type are unclear: this chapter fills this gap with an analysis of Tanzania. Considering this objective, VMLs are explored as a broader marketing channel involving not only contract farming schemes, but also other forms of vertical integration between agricultural production, the AP industry, and crop output markets. The VML category is sub-divided into two groups: the first, farmers' cooperatives, traders, and informal connections with small local businesses; the second, CF schemes and other direct sales to AP firms and more formalized businesses.

The empirical strategy addresses two main challenges: (1) selection bias and (2) unobserved heterogeneity. First, farmers with specific characteristics self-select into VMLs, and second, unobserved characteristics might drive estimates. The identification strategy is designed to address endogeneity by controlling for observed heterogeneity using propensity score matching, thus reducing selection bias. Then, I control for unobserved time-invariant heterogeneity using the panel data estimators on a balanced panel with three merged NPS-LSMS waves, which is an advantage over cross-sectional models. Despite my best efforts, it is not possible to fully control for time-varying unobserved covariates.

The main result is that VML participation leads to higher income for both VML groups and has significant effects for all outcome variables in the processors VML group. However, the significance of the effects for other outcome variables on the cooperatives VML group is reduced, with significant effects only on the benchmark OLS model. Moreover, subjective well-being presents statistical significance only in the FD model. A robustness check is conducted by estimating the ATTs using PSM that largely corroborate the panel results, however I caveat two endogeneities that limit the causality claim of the methodology: (1) possible omitted variable bias cannot be ruled out despite the set of control variables adopted; (2) I am not able to control for unobserved time-variant heterogeneity, although the shock variables absorb a significant share of it. These factors limit the causality claim of the methodology, so I prefer to interpret the results as associations.

I contribute to the policy debate by showing the positive effects of marketing channels on rural poverty in Tanzania (Ponte & Brockington, 2020) and the transformation of agricultural supply chains that it can bring in LDCs (Reardon et al., 2009). There is ample scope for LDCs governments to catalyze public policies on these topics especially improving agricultural market linkages in order to reduce transaction costs via better market access. For instance, collateralized credit, crop price support, the extension of social security to rural areas, harvest insurance, and clearer land tenure, all should be modified considering the specificities of SSFs. It is also noteworthy that extension services, such as agricultural best-practices training and improved seed distribution, could improve yields and agricultural productivity in a staple crop, such as maize, which is key for SSFs to first improve their food security and then increase their income via higher commercialization.

Overall, I conclude that a higher degree of market integration is strongly associated with improved welfare results for SSFs. Although I cannot claim that the results have a clear external validity for other regions, the results are representative of Tanzania, given the nature of the secondary data employed. Considering the characteristics of Tanzania's small-scale agricultural sector as similar to the African average, the results found are likely to be typical of those in similar countries. Nevertheless, follow-up research in various contexts would be useful to more precisely ascertain the magnitude of the effects and the specific pathways linking rural household welfare and VMLs with AP firms, local traders and vendors as well as CF schemes.

CHAPTER 3

Does contract farming improve rural welfare? Linkages between processing firms and Tanzanian cashew farmers[29]

3.1 Introduction

Dietary habits in developed countries are shifting toward healthier and more sustainable products (Gouel & Guimbard, 2019). This ongoing event favors the expansion of cashew nuts exports from tropical countries, a market that has grown sharply in the last decades. Following India and Vietnam, SSA is a pivotal producer of raw cashew nuts (RCNs). However, only approximately 10% of the regional output is processed locally, thereby siphoning off value from local producers and processors and dislocating African countries to the margins of this value chain (ANSAF, 2022). Consequently, the processing hubs and developed countries accrue the largest share of the increasing profits of the agro-processing (AP) industry by adding further value to the raw material through roasting, packing and, marketing, where the bulk of the value addition is concentrated (Tessmann, 2020). Cashew is a primary commodity, and thus, it belongs to a buyer-driven GVC, with low barriers to entry and where buyers determine producers' access to consumer markets (Gereffi et al., 2001).

Tanzania, which is the focus of this study, depends on the export of primary commodities to earn foreign currency and has low scores in economic complexity (Simoes & Hidalgo, 2012). Its most important cash crop is cashew nuts, with an export value amounting to USD 353.1 million in 2019, making them an important source of foreign exchange earnings and the mainstay of the livelihoods of over 700,000 households (Government of Tanzania, 2021). The

[29] **Publication status:** Bueno Rezende de Castro, A., Kornher, L., Magalhães de Oliveira, G., Rugaimukamu, K. (2023): Under review for publication as a journal article.

expansion of local processing of RCNs in Tanzania could lead to significantly higher demand for local production and it is estimated that the cashew nuts price can increase up to 10-fold from its raw state up until the pre-roasted cashew kernel (Government of Tanzania, 2021), thereby creating significant local value addition. Hence, Tanzania is an interesting case because: (1) it is one of the leading world producers of RCNs but with limited local capacity to process the raw material and market its final product; (2) the significant changes that the local cashew nuts sector has undergone since the partial market liberalization of the 1980s-1990s could be instructive for other transition countries; (3) it has a revealed comparative advantage (RCA) in the production of RCNs, but it underperforms in processing the raw material, which is a similar situation to other LDCs in SSA.

The prevailing marketing system for cashew nuts in Tanzania is based on an auction system that originated from the colonial past – similar to other SSA producers – that was partially liberalized since the 1980s-1990s (Kilama, 2013). In theory, the auction system reduces the pervasive problem of post-harvest losses (FAO et al., 2022) and has the advantage of increasing producer surplus by matching prices with the buyer’s willingness-to-pay. However, in reality, the auction system has been criticized as opaque and possibly corrupted, particularly in the licensing phase (FAO, 2015). Moreover, the auction system is not conducive to attracting more buyers because it introduces quality and quantity uncertainties for APs, lowering the prices received by farmers by reducing competition on the buying side. Contract farming (CF) offers opportunities for AP firms to offer higher prices to producers in exchange for lower market uncertainty. Thus, CF functions as partial insurance for SSFs (Bellemare et al., 2021) and provides a steady supply of raw materials to the industry. Although several studies have analyzed CF effects worldwide (Bellemare & Bloem, 2018), this chapter provides the first empirical evidence for the case of cashew nut farming in Tanzania.

This study investigates whether CF arrangements between SSFs and processors increase farmers' welfare compared to the standard government-run auction system, thus it focuses on the relationship between producers and AP firms, with a sub-focus on the opportunities for industrial upgrading and modernization of the cashew nuts' value chain in Tanzania via increased linkages to GVCs. I accomplish this objective by looking at the agricultural and industrial producers' side analyzing primary data in conjunction with supporting sectorial secondary data. Furthermore, a qualitative survey identifies a mediation channel of this effect: lower market and price risk. Studying the cashew nuts sector is essential as it is the linchpin of livelihoods in rural Tanzania, particularly in the south. Moreover, CF schemes in this value chain are a relevant study object because they can serve as means to improve rural development and increase rural value addition in producing countries (Meemken & Bellemare, 2020), which can stimulate the AP sector in LDCs.

Hitherto, the literature that established causal links of CF participation focused more on income, consumption, or production related variables (Bellemare & Bloem, 2018; Cahyadi & Waibel, 2016; Rao et al., 2012; Ruml et al., 2022; Ruml & Qaim, 2021a; Ton et al., 2018; Wang et al., 2014), leaving a gap of analyses of effects on food security (Bellemare & Lim, 2018; Bellemare & Novak, 2017; Mishra et al., 2018; Soullier & Moustier, 2018). Therefore, the present study fills a gap in the literature with a wider set of household welfare variables: I analyze the effect of CF on RCN output and household food security. This study contributes to the empirical literature by filling two gaps: (1) the lack of empirical studies analyzing whether CF arrangements in the cashew nuts sector improve farmers' welfare (Danso-Abbeam et al., 2022; Dubbert, 2019; Miassi & Dossa, 2019), particularly in Tanzania, and with auctions as the control group, rather than the frequently used spot market channel; (2) it adopts a mixed methods approach along with a joint market structure analysis using primary data from cashew nuts producers (SSFs) and processors, including CF determinants and sectoral challenges.

Using survey data of 339 farmers collected in 2022, I employed two empirical designs in this impact assessment: (1) a selection on observables by employing matching techniques, namely PSM and the IPWRA (Cahyadi & Waibel, 2016; Caliendo & Kopeinig, 2008; Wooldridge, 2007); and (2) a selection on unobservables using an IV approach to minimize sample selection bias, while also controlling for reverse causality and unobserved confounders (Cerulli, 2015). Although it is not feasible to use a fully randomized sampling strategy, stratification by treatment status at the village level ensures consistency. The main result shows that CF participation increases SSFs' output and decreases food insecurity by 0.13 and 0.40 standard deviations, respectively, compared to the auction marketing channel.

The rest of this chapter is structured as follows. After this introduction, Section 3.2 explains the cashew nuts market structure in Tanzania and presents statistics on the sector. Section 3.3 discusses the empirical strategy and the methodology adopted in primary data collection and Section 3.4 analyzes and discusses the results. Section 3.5 concludes the study and provides actionable and evidence-based policy recommendations.

3.2 Background and conceptual framework

3.2.1 The cashew nuts sector in Tanzania

The history of cashew nuts in developing countries is marked by varying degrees of state intervention according to different countries and times. In addition to political motivations, there are two explanations for this pattern. First, perennial tree crops such as cashew nuts have a relatively long maturation period (3-10 years), which makes them more vulnerable to international price volatility, thereby rendering producers unable to plan and respond quickly to price signals (Achterbosch et al., 2014). Second, cashew nuts are usually a large earner of hard currency, which is often scarce in LDCs (Huellen & Abubakar, 2021). In the post-independence period, the GoT built a processing capacity of 116,000 tons per year, which is still larger than the country's current capacity (Kilama, 2013). Following the demise of socialism in Tanzania in the late 1980s and the ensuing structural adjustment reforms, the sector opened to private investments and adopted a market-based regulation.

Cashew nuts has always played an outsized historical role in the domestic economy, however the sector started dwindling in the mid-1970s. One potential explanation is a fungi disease called powdery mildew that greatly reduced the trees' output and was not appropriately controlled following the villagization process[30] (Ngatunga et al., 2003). However, the reversal of fortune of the cashew nuts sector was not only due to an ill-conceived response to plant diseases, because failed policies also played important roles. During the socialist-oriented Nyerere era, the purchase of RCNs from farmers was centralized in a parastatal regulatory and processing firm (Cashewnuts Board of Tanzania, CBT) that re-invested the proceeds in agricultural extension services and local infrastructure, thus creating positive spillovers. The liberalization of the sector ended this scheme and shifted the burden, without an organized bureaucratic and political framework, to the central government budget (Kilama, 2013). Subsequently, the CBT became a board that regulates the sector but does not engage directly in production and marketing, except for market stabilization via purchases of excess supplies, a transition that also occurred in the vast majority of SSA countries (Swinnen & Maertens, 2007).

Although the cashew nut production has expanded with the liberalization, the same has not happened with the processing industry (Kilama, 2013). Since independence in 1961, both the Tanzanian output of RCNs and its monetary value have seen many ups and downs, hitting a nadir in 1990, and is now about 4 times larger (see Figure 3.1). Nowadays, cashew nuts represent 5.4% of Tanzanian exports, the second largest ticket after gold, with the country accounting for 3.1% of the global market, making it the seventh largest exporter[31].

[30] As part of the socialist Ujamaa policy, many farmers were compulsorily moved to new development villages.
[31] Data obtained from the FAOSTAT.

Consequently, it scores well in RCA[32]: together with other raw commodities, cashew nuts are placed in the fifth position in Tanzania's list of most specialized exported products (Simoes & Hidalgo, 2012).

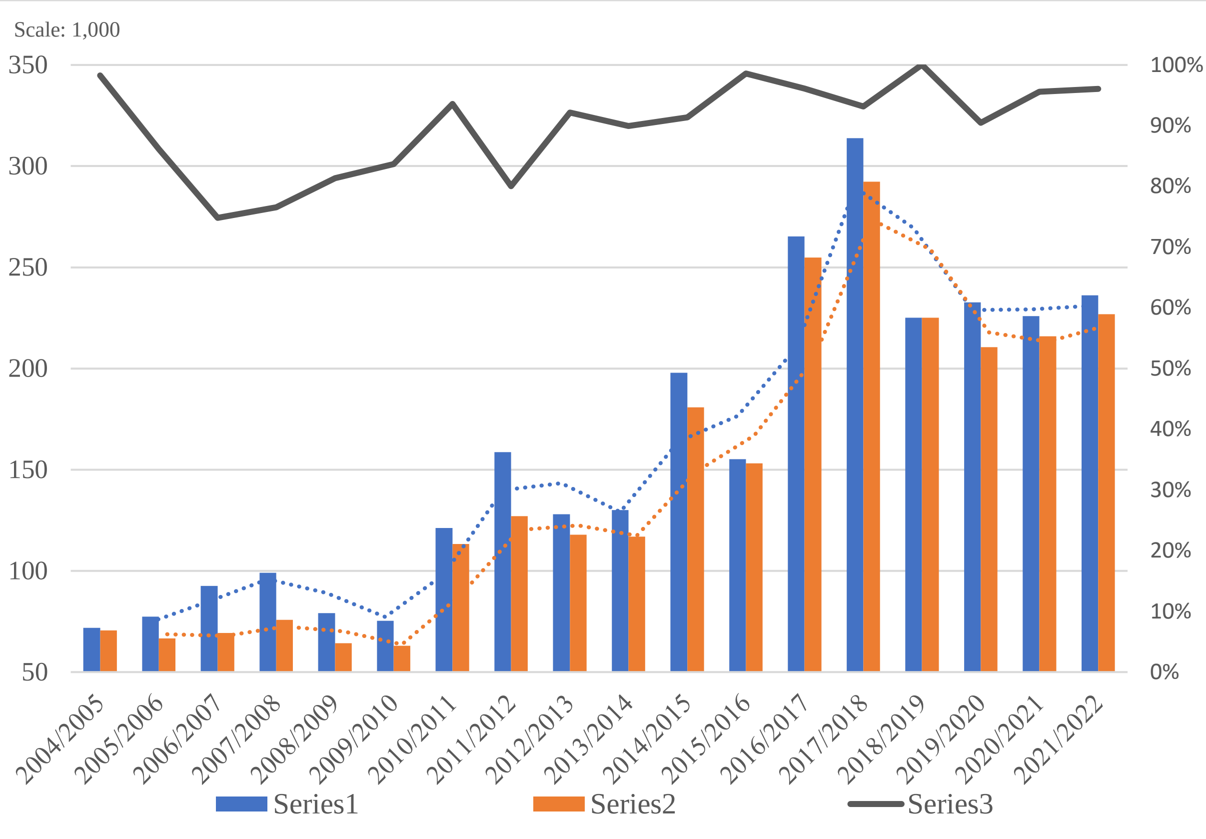


Figure 3.1: Recent trend of the cashew nuts sector in Tanzania

Notes: Left axis: 1,000 MT RCN output. Right axis: % of RCN exported. Source: Author's own elaboration based on CBT data obtained in the fieldwork.

The crop per se is categorized as a tropical medium-value perennial crop requiring some degree of processing before final consumption – other examples in the same category are coffee, cocoa, and palm oil (Feyaerts et al., 2020). The value chain is characterized by a sharp separation between production and processing, which also tend to be activities geographically separated: value addition is generally conducted outside of the farm and overseas – in Asia, at the initial processing stages, and in high-income countries (HICs), at the final stage. These processes are lightly capital-intensive that, although they were once located in some producing regions, they ended up being dominated by middle-income countries with more capital and medium-skilled

[32] Using data from the FAOSTAT, I calculate a revealed comparative advantage in RCNs equal to 108.

labor available. Processing firms and cooperatives located in the Global South remain subordinately inserted into GVCs (especially buyer-driven value chains), as reflected in the very low exports of processed RCNs from SSA countries. This case translates into a small degree of value addition to RCNs and lack of access to foreign investments and funding, therefore, local firms have limited opportunities to functionally upgrade to higher value-added segments in the value chain (Gereffi et al., 2005; Huellen & Abubakar, 2021).

The RCN value chain in Tanzania has a bipolar governance structure similar to the Ivorian case (Tessmann, 2018): trader-driver in the segment between SSFs and Asian processors, whereas buyer-driven in the segment between local processors and retailers in HICs. In this regard, although there are many possibilities for cashew nuts products and functional upgrading (e.g., shelling, packing, and marketing), adherence to global standards and production systems – given that the bulk of the final product is consumed in HICs – is a binding constraint. Food quality and sanitation standards from HICs have always been stringent; thus, Tanzania still needs to build institutional frameworks to comply with these demands.

The Tanzanian cashew nuts sector has an intermediate degree of consolidation as most smallholder farmers and cooperatives produce through horizontal integration. Part of the little processing done in the country is conducted close to – but not directly at – the cashew tree fields in small- and medium-scale firms. However, the bulk is the responsibility of large firms. Following the model of other tropical commodities, domestic consumption is minimal, production is almost completely oriented to export markets, and the processing industry is relatively concentrated. Considering a total of 45 firms registered with the CBT, the 10 largest firms account for 60% of the domestic processing; and 90% of Tanzania’s output is exported unprocessed[33]. The market is defined as an oligopsony, thereby weakening the bargaining capacity of producers, particularly smaller ones, that mostly rely on auctions (Kilama, 2013;

[33] Figures obtained by the author in interviews with the CBT in March-May 2022.

Porto et al., 2011). Consequently, an over-reliance on exporting RCN instead of processed cashew nuts exposes the country to a less competitive market and a loss in terms of trade (McMillan et al., 2002).

The bulk of the production of cashew nuts in Tanzania comes from the Mtwara and Lindi regions in southeastern Tanzania, close to the border with Mozambique, another leading producer of cashew nuts, and from the Pwani region, encircling Dar es Salaam, the economic capital. Due to the geographical concentration of production in these regions, the sector benefits from agglomeration effects, that partially reduce transport and inputs search costs, while increasing the supply of middle-skilled labor and the possibilities for accumulating knowledge and know-how between firms and workers (Glaeser, 2010; Venables, 2008). Figure 3.1 shows the production and export trends of the sector.

Moreover, this sector has an inherent geographical advantage: together with Mozambique, Tanzania is the only major RCN exporter located in the Southern Hemisphere, therefore, the harvest occurs at the opposite time compared to the large Northern producers. This fact confers the country an advantage as it markets its produce when there is a cyclical shortage, which can help farmers command higher prices. There is, however, a need for government policies to better explore this advantage.

3.2.2 Governmental policy in the cashew nuts sector in Tanzania

In order to illustrate the importance of the crop for the GoT, I quote the Director-General of the CBT[34]:

> "More research at the level of CBT, universities, and TARI-Naliendele needs to be conducted to determine why cashew nuts do not benefit the farmers as much as they should. This a true puzzle, and I think the solution is for Tanzania to finally move away from a reliance on exports of RCNs and, instead, process them in the country. Only then we will accrue most of the benefits through domestic

[34] Source: Interview conducted by the author during fieldwork in March 2022 with Francis Alfred.

value-addition".

The government intervention in the Tanzanian cashew nuts sector is conducted through two main regulations: Act No. 21 of 1984 and Act No. 18 of 2009. The first one was introduced at the twilight of Nyerere's administration to create the Cashewnut Marketing Board tasked with regulation, licensing, marketing, and coordination of all cashew nuts activities. The second act replaces the first and rebrands the governmental parastatal as CBT which, among others, now authorizes CF but only under its purview, introducing some flexibility to producers (Cashewnut Industry Act, 2009). Nowadays, both cashew nuts agricultural production and its processing industry figure prominently in the policy agenda of the GoT. Notwithstanding its strategic importance, there does not seem to be in place a coherent strategy for industrial development that would allow Tanzania to functionally upgrade in the value chain and, more importantly, generate employment in rural and peri-urban areas – where most of the processing is located.

In addition to the aforementioned regulatory acts, the Agricultural Sector Development Strategy states as objective the strengthening of both the production and domestic processing of traditional cash crops through export processing zones and regular consultation with private sector stakeholders, albeit it neither goes into further detail, nor take stock of the success of past policies and measures (Government of Tanzania, 2015). The most recent National Five-Year Development Plan places cashew nuts as one of the priority crops for a series of high-level government interventions: (1) increase the use of ICTs on production and processing alongside modern crop management systems; (2) enhance investments on research and development of high-yield crops; (3) improve irrigation systems and expand sustainable water and land use management (Government of Tanzania, 2021).

Despite the clear articulation of the national cashew strategy in policy documents, the implementation falls short of the objectives set by the GoT, thus preventing the sector from quickly progressing. One important shortfall is that the sets of aforementioned policies lack an actionable checklist of means as well as concrete and quantifiable goals. Almost all key stakeholders interviewed indicated that there are pervasive issues with not only the implementation of public policies for the sector but also with faulty coordination of GoT actors. Last but not least, cashew nuts also have relevant political importance due to the electoral power of the constituencies in the main producing regions. The interviews with key experts and firms suggest that the GoT tends to favor farmers – a sizable group of voters – when there are conflicts of interests with processors on policy making.

3.2.3 Conceptual framework

Tanzania is transitioning from small-scale semi-commercialized agriculture to slightly larger scale commercialization featuring some nascent new forms of market linkages with processing firms and international buyers. This pattern is also reflected in the cashew nuts sector as its local supply chain is based on agricultural cooperatives[35], which sell RCN either at CBT auctions to bulk buyers (often traders), or directly to AP firms with which they have contracts, which are generally, increasing market risk and reducing the scope for repeated transactions (Fafchamps & Minten, 2001).

A total of four main actors are identified in the cashew nuts value chain as shown in Figure 3.2: farmers and farmers' cooperatives, who are assumed to be risk-averse, AP firms, traders (or middlepersons), and the government. I also assume that there are two broad types of RCN quality that drive prices: whole or broken. The raw material is generally more homogeneous than in other agricultural value chains, which is an important difference that leads to relatively fewer refusals by AP firms. By contrast, for example, in the dairy sector, the milk quality varies much more, thereby determining a large share of the price that farmers and cooperatives receive (van Campenhout et al., 2021). Despite this difference, and in part due to the historical

[35] Mostly referred to in Tanzania as "agricultural marketing cooperative societies", or AMCOS.

centralization, the sector is organized around cooperatives.

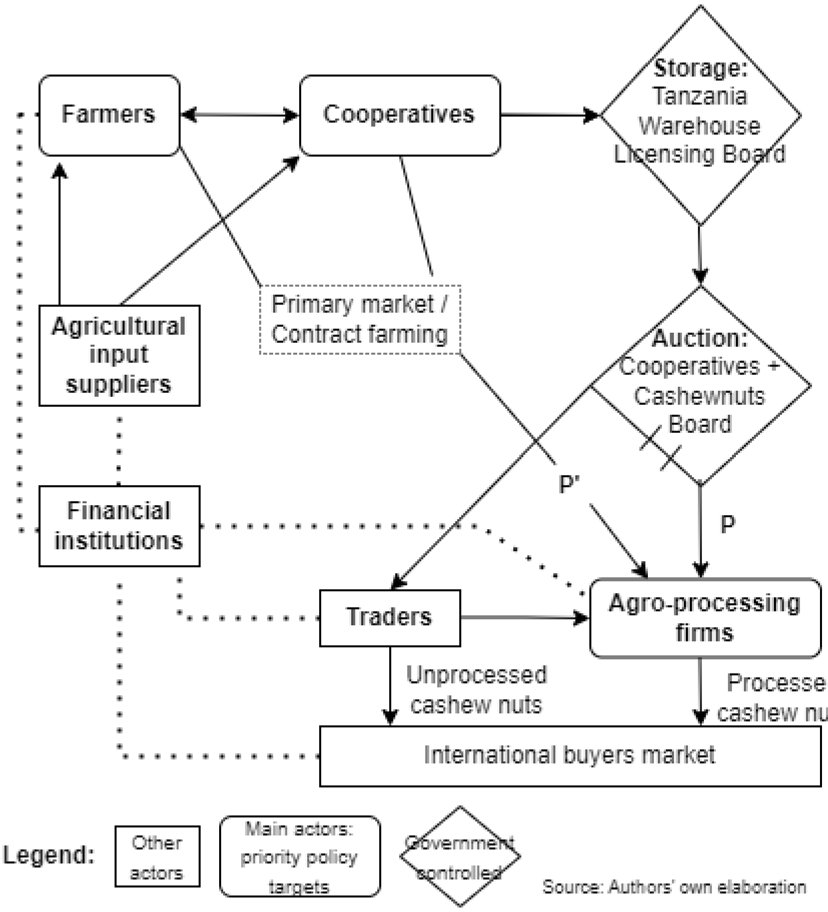


Figure 3.2: Structure of the cashew nuts sector in Tanzania

Notes: P = cashew auction price, P' = cashew CF price. P' > P.

The sector operates via a market system with government regulations and compliance with CBT guidelines[36] with regard to auctions. Among those pertaining to the purchase of RCN, the most important is to prohibit buying outside of the warehouse receipt system (WRS), which consists of government-run storage facilities located close to production areas, thereby reducing transport costs and, thereafter, storage costs. The transportation of the RCNs follows a three-step sequence: (1) farmers deliver the majority of their output to primary societies – small associations of farmers in nearby villages, the backbone of the cooperative system –, which

[36] Two pieces of anecdotal evidence: (1) in 2013 riots broke out in the South when the auctions paid much less than what farmers expected per kg of RCNs; (2) in 2018 the central government deployed the military to purchase RCNs at a premium directly from farmers because they were not satisfied with traders' bid at the CBT auction and, thus, became, together with AP firms, politically disgruntled with the GoT. A subsequent temporary export ban on RCNs coupled with a temporary policy of a minimum indicative price led to an abrupt decline in the country's exports.

then aggregate the combined produce in collection centers; (2) all cashew nut farmers are mandated to become members of a cooperative, which then transports the produce from collection centers to the storage of the Tanzania Warehouse Licensing Board[37]; (3) from there, the produce is distributed to the auctions' sites, most commonly in the regions of Mtwara and Lindi. AMCOS also represent farmers in auctions, which are organized by the CBT; prices are set by market forces, but the CBT regulates auctions to ensure that farmers receive a fair price[38]. Farmers also receive subsidized inputs (mostly fungicides and pesticides) and partial advance payments that are paid with their produce posted as collateral in the WRS. Two actors are licensed to participate in the auctions: traders, who often represent foreign buyers and intermediate the export of RCN, and local processing firms, which buy RCNs as a raw material. The only exception to this controlled formal market is the recently launched primary market, wherein Tanzania-based processing firms can directly purchase RCN at the farm gate, thus avoiding competing for prices in the auctions, which is a common complaint from local firms[39]. Moreover, the mechanism aims to ensure that the domestic industry can have immediate access to enough raw materials, irrespective of market circumstances; hence, the primary market system works as CF.

In general, the literature documented that CF has many common benefits over spot market transactions and auctions: a higher degree of coordination with upstream value chain actors; more attention to produce quality and harvest timing; reduction of market risk (Adabe et al., 2019; Barrett et al., 2012; Bellemare & Bloem, 2018; Meemken & Bellemare, 2020). Furthermore, farmers working in cooperatives tend to fetch higher prices for the following reasons: (1) better bargaining power in sales negotiations by reducing the profit margins of both buyers and middlepersons (de Brauw & Bulte, 2021); (2) easier and cheaper access to

[37] With two fees involved: farmers pay Tsh 50 per kg of produce as AMCOS' transportation fee and pay between Tsh 25-60 per kg for storage costs of the RCNs in the governmental warehouses, depending on distances.
[38] This information is based on primary interviews with CBT managers, who do not disclose what it considers a "fair price", a concept that can be approximated to a floor price.
[39] However, the minimum bid at the auctions is still considered high (50 tons) for some smaller local firms.

mechanization, high-quality inputs, and technical training (B. Lin et al., 2022); (3) bulk purchase of agronomical inputs (Campenhout et al., 2021). These effects are analogously observed in CF schemes. In particular, in this chapter's sample, the prices fetched by CF participants are approximately 13% higher than those by control farmers.

This study's fieldwork identified only the marketing contract type where parties agree on RCN prices, quantities, and transportation arrangements, without provision of neither specific agronomic practices/techniques nor inputs and/or credit, e.g., resource-providing contracts. The appropriateness of analyzing a single type of contract is backed by evidence from a previous field experiment in Benin, which found that a simple contract resolving price risk is nearly as beneficial to SSFs as more complex resource-providing contracts (Arouna et al., 2021). The exclusive prevalence of marketing contracts is primarily due to the underdeveloped stage of CF in this sector in Tanzania, as it is a recent institutional innovation from the late 2000s (Kuzilwa et al., 2019). Further reasons are skepticism from policymakers based on political economy reasons[40] and high transaction costs for the negotiation, signing, and enforcement of contracts (Williamson, 1979). The latter is especially true in environments with a weak rule of law, and it is precisely this market failure, coupled with incomplete markets, that CF can alleviate by increasing trust between contracting parties, thereby potentially increasing SSFs' welfare (Maertens & Vande Velde, 2017). I hypothesize that participation in CF affects SSFs' welfare, particularly RCN output and food insecurity, by insuring them against market and price risk, which is reflected in the predominant contract type (Arouna et al., 2021; Bellemare et al., 2021). Previous study suggests that contract farmers tend to be more risk averse than independent farmers (Mishra et al., 2020).

Finally, for the purposes of this dissertation, participation in CF schemes is considered an

[40] Based on field interviews with GoT actors in March-May 2022, I found that they worry that processors might exploit farmers often exceeds perceived welfare benefits.

efficient solution for SSFs to reduce the considerably high transaction costs that they face in rural markets. This happens precisely because CF partially insures farmers against the high market and price risk that they face in spot markets and subsistence production. In this sense, this dissertation does not directly quantify the size of transaction costs but rather it estimates the effects of CF participation on SSFs income and food security. The main contribution of this chapter, which is the estimation of these effects, is then an approximation of the magnitude of the toll that transaction costs exact in the Tanzanian cashew nuts sector.

3.3 Methodology

The empirical analysis presented in this section is divided into two parts: (1) a description of the primary data collection and sampling strategy, which used semi-structured quantitative and qualitative interviews with farmers, AP firms, and key stakeholders (policymakers) – due to time and resource limitations, I focused on the three main actors in the value chain; (2) an explanation of the empirical strategy to estimate the effects of CF participation on farmers' household welfare complemented by qualitative analyses.

3.3.1 Data collection

This study relies mainly on primary data collected in Tanzania between March and May 2022 by the main author of this study together with three trained enumerators; and then supplemented by secondary sources for descriptive statistics. The primary data were collected downstream with cashew nuts farmers using a farm household survey and upstream with AP firms – geographical locations are shown in Figure 3.3, covering the most prominent cashew nuts-growing regions in the country. The interviews were not recorded, but the answers were inserted into tablets that were pre-programmed with the pre-tested survey questionnaire translated into Swahili, which is common in this type of research (Rutakumwa et al., 2020). Qualitative questions were also included in the survey in order to have a deeper understanding of the quantitative data, e.g., regarding potential transmission mechanisms.

Through interviews with the farmers, we collected data on household characteristics, agricultural production, income generating sources, connection to processing firms (unavailable in secondary data sources), and socioeconomic shocks. All quantitative variables and indicators are objective measures to avoid self-reported bias, and Table 3.1 presents the main variables. I include a series of control variables that are expected to influence the outcome variables and selection into CF in order to minimize omitted variable bias that can affect CF participation. The main variable of interest is a dummy capturing whether households participate or not in CF and I make two caveats: first, in this chapter's context, measuring the intensive margin of CF was not practical because cashew nuts are a cash crop that households rarely consume themselves; second, I can safely rule out measurement errors because we interviewed farmers selected from cooperatives' rosters and there is no incentive or social stigma for why farmers would not give an accurate response to their CF participation status (Bellemare et al., 2021). Household food insecurity access scale (HFIAS) as an indicator of food insecurity follows the canonical definition that is synonymous with the quantity and quality of food ingested (Coates et al., 2007).

To inform the empirical analysis, we also conducted qualitative semi-structured interviews with 13 processing firms and 23 key stakeholders in the cashew nuts sector using semi-structured questionnaires. The apparently small sample size of processors reflects a substantial proportion of the universe: from a total of 45 cashew nuts processing firms that have installed capacity in Tanzania, only 37 are currently in operation[41], and we interviewed 13 of them (35%). The interviews with the processing firms collected data on output, innovation and technology

[41] Data were obtained by the author in meetings with CBT officers in March-May 2022.

adoption, marketing strategies, and connection to GVCs.

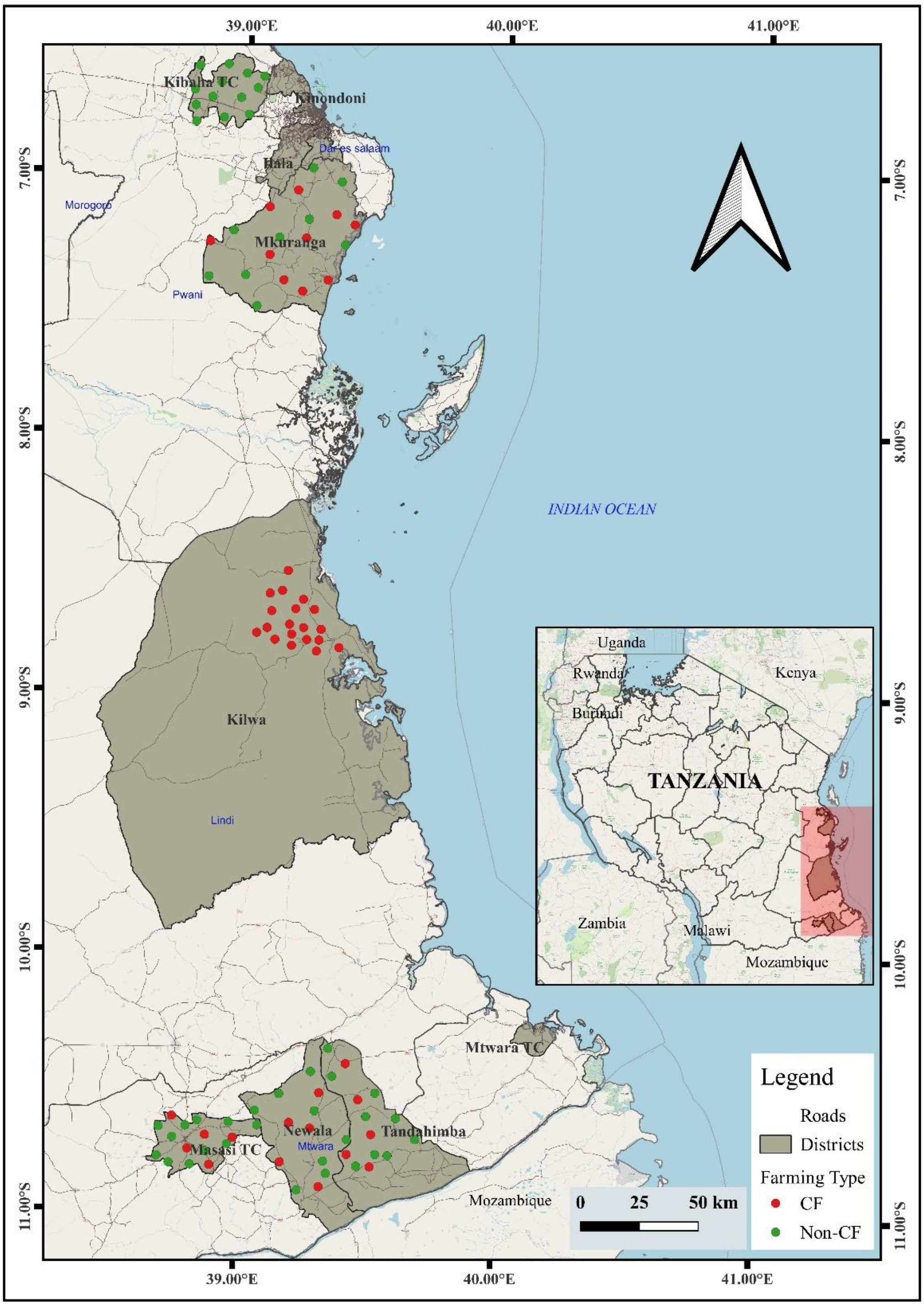


Figure 3.3: Sampling locations of cashew nuts producers and processors in Tanzania

Source: Author's own elaboration using GPS coordinates from fieldwork primary data layered onto a Tanzania shape file on *QGIS*.

Table 3.1: Construction and definition of the variables in the farmers' household questionnaire

Dependent/outcome variables

Cashew nuts output	Quantity of raw cashew nuts harvested in the last 12 months, in kg

Food insecurity	Household food insecurity access scale (HFIAS)

Independent/control variables

Socioeconomic characteristics

Household size	Number of household members
Age	Age of household head, in years
Years of schooling	Years of formal education of household head
Gender	1 if the household head is a female, 0 if a male
Access to credit	1 if the household had access to credit, 0 if otherwise
Migrant	1 if the household head is not originally from the region where the farm is located, 0 if otherwise
Agricultural characteristics	
Contract farming	1 if the household sold cashew nuts to processing firms in the last 12 months, 0 if otherwise
Farm size	Farm area dedicated to cashew nuts, in acres
Family labor	Number of workers/hours that the household used from the family in the last 12 months
Hired labor	Number of workers/hours that the household hired in the last 12 months
Pesticides	1 if the household used pesticides, herbicides, and/or fungicides in the last 12 months, 0 if otherwise
Soil	1 if the quality of the soil of the farm plot is good or excellent, 0 if otherwise
Steep	1 if the soil of the farm plot is located in steep terrain, 0 if otherwise
Training	1 if the household head has participated in agricultural training, 0 if otherwise
Regional characteristics	
District	District term

3.3.2 Sampling strategy

I used a multistage random sampling procedure to select the participants of the interviews based on their relevance in the value chain, in the case of the processors, and their representativeness, in the case of the farmers. Regarding the latter, the sampling strategy seeks to minimize potential biases between contract and traditional farmers, which comprise two strata: farmers that sell only in auctions (via their cooperatives), and those who have CF schemes with AP firms – in the latter case, we sampled from three firms, with each sourcing from one of the three regions. In this regard, the institutional and market context of the study area minimizes some issues of selection bias because the nonparticipation of farmers in CF is driven by not only unobservable characteristics (e.g., innate ability and motivation) but also due to supply-side constraints: only a handful of CF schemes ongoing in the Tanzanian cashew nuts sector exist, the most important of which are captured in this study. However, it is unknown if all farmers would choose to participate in case the schemes were available.

In the first stage, I selected the targeted regions where the bulk of Tanzanian cashews are produced. Second, I selected districts and villages that are representative of cashew production and where CF also takes place. Thereafter, we met leaders from eight different cooperatives, including CBT and local government officers to obtain lists of contacts of members from which individual farmers were randomly sampled and interviewed to form the cross-section sample. The total sample size consists of 339 farmers[42], from which 117 engage in CF and 222 do not – these groups are mutually exclusive, thereby complying with the stable unit treatment value assumption. I also used non-proportional random sampling to select farmers in each of the two strata such that the sample could have a sufficient number of contract farmers – because there tend to be fewer contract farmers in villages.

3.3.3 Empirical specification

I estimate the following main model:

$$Y_i = \beta_0 + \beta_1 CF_i + \beta_2 \boldsymbol{X_i} + \varepsilon_v, \tag{3.1}$$

where *i* indexes each farming household, *Y* denotes the outcome variables (all estimated separately to avoid simultaneity bias), $\boldsymbol{X_i}$ is a vector of socioeconomic, agronomic, and regional controls. The error term ε is clustered at the village level for three reasons: (1) this term was the unit of randomization of the sample selection design; (2) the treatment and/or sampling assignment could be correlated within each cluster; (3) processors choose CF participants at the village level (Abadie et al., 2022). The parameter of interest is β_1, which estimates the effect of CF, captured by the binary variable *CF*, which encompasses SSFs who sold their produce directly to processors.

[42] The power calculation is shown in Appendix 7.2.1.

I use OLS as a regression benchmark and a Probit model to assess the association between CF participation in the Tanzanian cashew nuts sector and farmers' household welfare indicators. As these β_1 estimators are prone to sample selection bias, I adopt an IV approach to more accurately estimate the impact of CF participation on farmers' welfare outcomes whereby the instrument better controls for unobservable variables, reverse causality, and self-selection into CF (Wooldridge, 2010). In addition, I make use of two other closely related estimators, PSM and IPWRA, for robustness purposes. The purpose of the combination of these estimation methods is to minimize endogeneities related to reverse causality and selection on unobservables.

3.3.4 Selection on observables

PSM and IPWRA use different mechanisms to calculate the average treatment effect (ATE) on the treated (ATT), given by the following:

$$ATT = E\left(Y_{i,CF} - Y_{i,A} \middle| CF_i = 1\right) = E\left(Y_{i,CF} \middle| CF_i = 1\right) - E\left(Y_{i,A} \middle| CF_i = 1\right), \tag{3.2}$$

where the variables and indexes are the same as in Equation 3.1, in addition to *A*, which stands for "auction", that is, the traditional marketing channel (control group). As this study is observational given that perfect sample randomization was not feasible, the challenge of the identification strategy is to estimate the counterfactual scenario for $Y_{i,A}$, i.e., when $CF_i = 1$. This is not observable, since, in the data set, farmers having contractual arrangements do not utilize the auction marketing channel, hence I cannot observe outcomes for treated observations as if they were not treated and vice-versa.

The ATTs are computed with propensity scores representing the probability of CF participation, given by the following:

$$p(X') = \Pr(CF_i|\boldsymbol{X_i}) = E(CF_i|\boldsymbol{X_i}), \tag{3.3}$$

where $\boldsymbol{X_i}$ is the same vector of covariates from Equation 3.1 and $p(\boldsymbol{X_i})$ is the propensity score.

PSM is performed by two matching algorithms, i.e., radius and kernel. The former attributes weights based on the matched covariates' distance between each observation in the control group and the counterfactual estimated for the treatment group; and the latter retains more observations from the control group (Caliendo & Kopeinig, 2008).

The main advantage of the IPWRA estimator over PSM is that it goes beyond ATTs whereby it predicts non-observable counterfactual means adjusted for the regression's covariates, thereby computing the ATE by taking the difference between treated households (CF participants) and their counterfactual non-treated means (auction farmers), given by the following:

$$ATE = E\left(Y_{i,CF} \middle| CF_i = 1,0\right) - E\left(Y_{i,A} \middle| CF_i = 1,0\right). \tag{3.4}$$

Given that PSM is still prone to bias from unobserved confounders (Imbens, 2004), the IPWRA's doubly robust property is desirable because it needs only one of the two estimated models (i.e., treatment and outcome) to be correctly specified, without making assumptions about the functional form, to generate consistent estimates (Wooldridge, 2007).

The IPWRA estimation process consists of four steps: (1) first, estimate the parameters of the selection equation to predict treatment status and compute the inverse of the treatment probability (IPWs); (2) second, estimate separate regression outcomes, one for treated (CF farmers) and one for untreated households; (3) third, compute the means of both groups using the IPWs; (4) finally, obtain the ATEs by calculating the differences between the two weighted means. A variation of the last step yields the ATTs if the calculation is restricted to the treatment group. The weighting uses 1 for the treated and uses the inverse of the probability of being in the control group for the non-treated, whose weights are given by the following:

$$(3)\ \widehat{w_i} = CF_i + (1 - CF_i) \times {\hat{p}(\boldsymbol{X_i})}\big/{(1 - \hat{p}(\boldsymbol{X_i}))}. \tag{3.5}$$

3.3.5 Selection on unobservables

As processors decide the cooperative and village in which they will select farmers, the OLS estimator of the CF effect may be endogenous, thus leading to a biased estimate of β_1 in Equation 3.1. The PSM and IPWRA estimators address this only based on observable differences between the treatment and control groups. I employ an IV approach to further address endogeneities related to reverse causality and correlation of unobservables with the error term, wherein predictions of the treatment variable using an instrument from a first-stage regression are included in a second-stage outcome regression to minimize correlation with the error term (Imbens & Wooldridge, 2009; Wooldridge, 2010). The IV two-stage least-squares (2SLS) estimator reduces bias compared to OLS or the aforementioned matching methods, particularly in overidentified models, as is our case (Davidson & MacKinnon, 2007). In order to implement the IV-2SLS approach, Equation 3.1 is modified as follows:

$$CF_i = \alpha_0 + \alpha_1 \boldsymbol{Z_i} + \alpha_2 \boldsymbol{X_i} + u_v, \tag{3.6}$$

where CF participation is explained by the same vector of controls in addition to a vector $\mathbf{Z_i}$ with the excluded instruments, and the error term u. In the IV second stage I replace the variable of interest (CF) by the predicted values of the instrument from the first stage:

$$Y_i = \beta'_0 + \beta'_1 \boldsymbol{Z}_i + \beta'_2 \boldsymbol{X_i} + \varepsilon'_v. \tag{3.7}$$

I use two main classes of instruments for CF participation: the number of CF participants in a certain radius around each farmer and the distance from each farmer to the nearest CF participant (in Km) – I also include the latter IV squared to account for non-linear effects (Fischer & Qaim, 2012a). The baseline model using the IV estimator adopts radiuses of 10 and 100 km; additionally, I check for robustness by testing three other models with the following radiuses in Kms: 10-200, 25-100, 25-200 (see Tables 3A.2-3A.5). These IVs are motivated by studies that used distance-based instruments in the context of CF or similar vertical integration schemes (Chege et al., 2015; Dubbert, 2019; Miyata et al., 2009; Ruml & Qaim, 2020; Tambo et al., 2020). Table 3A.1 presents the correlation matrix between the instruments and the treatment variable showing that the instruments are relevant in influencing CF participation.

The instrument choice is based on the literature on technology adoption and peer learning, which posits that these instruments can capture social network effects (Maertens & Barrett, 2013; Michelson, 2017; Sellare et al., 2020). I assume that SSFs are risk-averse, therefore the rationale is that they are hesitant to participate in CF schemes compared to the default market channel (CBT auctions). Hence, they prefer to observe their peers' behavior first and then choose to adopt it if they believe that it will benefit them and have lower price and market risk compared to the auctions (Ma & Abdulai, 2017). As previously discussed, CF in the Tanzanian cashew nut sector is still incipient and, in the sample, the processors offer the CF schemes after receiving approval from the CBT and from cooperative leaders; thereafter, individual farmers are free to decide whether to join.

Furthermore, the unconfoundedness assumption holds because there is no reason to believe that the instruments share unmeasured common causes with neither the outcomes nor the treatment, further corroborating the results of the endogeneity tests (see Table 3A.6). In rural East Africa (including Tanzania), farmers tend to live on the land their families cultivate, resulting in a scattered geographical pattern within each village (Fischer & Qaim, 2012b; Lowder et al., 2016). First, farmers are unlikely to have any power in deciding who their neighbors are because plots are usually allocated via inheritance across farmer generations. Second, acquiring new governmental land leaseholds is difficult in that setting.

However, AP firms may choose to offer contract farming to SSFs close to their factory, meaning improved access to infrastructure. In Figure 3.3, apart from the Lindi region[43], no clear

[43] The results are robust to excluding observations from the Lindi region. These additional empirical exercises are available upon request.

geographical clustering of CF farmers is observed, which means that a counterfactual where processors chose neighboring farmers to contract because of structural similarities with respect to actual CF participants is likely. I perform an additional robustness check to test if the distance from each household to the nearest AP firm is associated with improvements in the outcome variables. This placebo test can suggest whether differences between treatment and control group are indeed the result of the treatment or occur by chance.

The results from this placebo test (see Table 3A.10) do not hint at unobserved correlation of the outcomes with the geographical location, as statistically significant effects were not found. This finding stresses the idea that the distance to AP firms does not affect RCN output and food security. Hence, while, in theory, farmers located near to processors are more likely to participate in CF due to lower transportation costs, SSFs from the sample do not have naturally higher levels of the outcomes here analyzed. Moreover, the three processing firms did not adopt a systematic practice of choosing farming households nearby. In conjunction with the IV tests, these robustness checks make the distance-related instruments plausibly exogenous.

The identifying assumption for this study is that the distance from each farmer to the nearest CF participant is correlated with CF participation, but not correlated with cashew nuts output and food insecurity in any other way than through CF participation. As aforementioned, the theoretical reasoning is that the presence of CF farmers in neighboring areas plays an important role in farmers participating in contract farming. Furthermore, I address some potential issues present in other empirical studies and those pointed out by Bellemare & Bloem (2018). First, I include a district term to control for regional effects such as better access to markets and/or information. Second, I control for agroecological conditions to account for the possibility that CF farmers may be located in areas with better soil quality. Finally, I adopt a set of household controls, including education, access to credit, and agricultural training, that could potentially confound the empirical design. Together, all these points strengthen the identification strategy, making it more plausible.

Lastly, I argue that the exclusion restriction holds because: (a) no clear pattern of geographical clustering separating contracting farmers from the control group exists; and (b) falsification tests employ separate OLS regressions of the instruments on the outcome variables with the full set of controls showing no statistically significant effect[44] (see Tables 3A.2-3A.5) (Di Falco et al., 2011). I perform the following additional statistical tests (see Table 3A.6): (1) large first-stage R-squared and significant F-statistics; (2) the Wu-Hausman test cannot reject the null hypothesis that the IVs are exogenous; (3) the overidentifying restriction test (i.e., four excluded instruments for one endogenous variable) results that the IVs are uncorrelated with the error term; (4) the Chi-squared statistics reject underidentification; (5) and the F-statistics ensure that the IVs pass the weak identification tests (Stock & Yogo, 2005; Wooldridge, 2010).

3.4 Results

3.4.1 Descriptive statistics

Cashew production is the main source of monetary income for all sampled households, thus indicating the decisive role the crop plays for these farmers. While this dependence might be harmful in the sense that the income risk is too concentrated on a single source, it is also true that this cash crop is produced in relatively poorer parts of Tanzania with few off-farm employment options (especially in the South) – on top of infrastructural issues hampering access to markets.

Table 3.2 presents a summary of the descriptive statistics showing the differences in the sample between CF households and those preferring the traditional marketing channels (auctions) in the Tanzanian cashew nuts sector; CF farmers have an RCN output that is approximately a third

[44] Moreover, the joint significance test of the instruments results in the Chi-square statistics show joint statistical significance of all instruments in determining CF participation but not on the outcomes.

statistically significantly higher than non-CF actors and also lower food insecurity scores. Furthermore, I highlight that CF farmers achieve this result with a statistically significant similar amount of land, which may be a result of a higher number of trees and a greater exposure to agricultural training – no statistically significant differences are observed between application of pesticides, fungicides, and herbicides, soil quality, and labor input. Given that the outcome variables are not normally distributed, nonparametric tests of the differences in the outcome variables' distributions are shown in Figure 3.4, indicating a left skewness – complementary boxplots are shown in Figure 3A.2. Notably, the treatment and control groups partially overlap in some parts of the distribution, thus presenting little systematic difference in the drivers of outcome variability, although, naturally, these figures only capture unconditional correlations. Finally, some of the control variables are statistically significantly correlated with the outcome variables and, thus, any estimation of the CF participation effect on the outcomes needs to be controlled for these characteristics[45].

Table 3.2: Descriptive statistics for outcome and control variables

	(1) Full sample	(2) CF	(3) Non-CF	(4) Test of difference
RCN output (kg)	1283.16	1521.87	1160.67	-361.21**
	(1329.41)	(1560.72)	(1178.58)	(0.02)
Food insecurity (HFIAS)	2.19	1.75	2.41	0.66***
	(1.16)	(0.97)	(1.18)	(0.00)
Food insecurity (HFIAS normalized)	0.19	0.13	0.23	0.10***
	(0.23)	(0.19)	(0.24)	(0.00)
HH size	5.03	5.31	4.88	-0.43*
	(2.01)	(2.16)	(1.92)	(0.06)
Age of HH head	53.86	51.51	55.07	3.56**
	(13.48)	(13.97)	(13.10)	(0.02)
Years of education of HH head	7.88	6.74	8.47	1.73***
	(3.10)	(3.27)	(2.84)	(0.00)
Gender of HH head (Female = 1)	0.17	0.09	0.22	0.13***
	(0.38)	(0.28)	(0.41)	(0.00)
Access to credit (0/1)	0.13	0.07	0.16	0.09**
	(0.34)	(0.25)	(0.37)	(0.01)
HH is not migrant (0/1)	1.19	1.07	1.26	0.19***
	(0.40)	(0.25)	(0.44)	(0.00)

45 Results are available upon request.

Farm land under cashew nuts (acres)	8.70	8.07	9.02	0.95
	(10.76)	(7.08)	(12.22)	(0.44)
Family labor (days per year)	73.76	71.98	74.67	2.68
	(77.00)	(69.36)	(80.77)	(0.76)
Hired labor (days per year)	27.06	24.83	28.20	3.37
	(39.37)	(34.08)	(41.85)	(0.45)
Application of herbicide/pesticide/fungicide (liters)	215.20	236.26	204.39	-31.86
	(264.96)	(267.61)	(263.52)	(0.29)
Total number of cashew nuts trees	275.39	394.21	214.41	-179.81***
	(427.00)	(630.54)	(249.25)	(0.00)
Soil quality good or excellent (0/1)	0.63	0.64	0.62	-0.02
	(0.48)	(0.48)	(0.49)	(0.74)
Soil steep (0/1)	0.07	0.04	0.08	0.04
	(0.25)	(0.20)	(0.27)	(0.20)
Received agricultural training (0/1)	0.53	0.72	0.44	-0.28***
	(0.50)	(0.45)	(0.50)	(0.00)
Observations	339	117	222	

Note: Standard deviation in parentheses. Significance levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Mean difference: t-test.

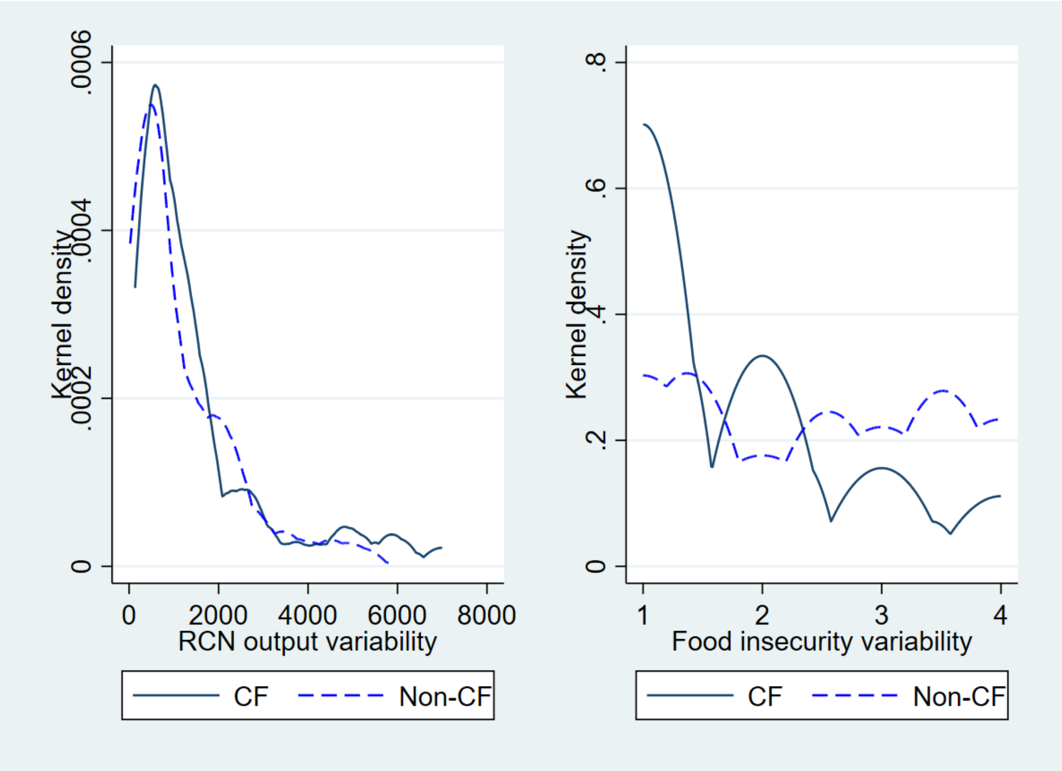


Figure 3.4: Kernel density estimates of the variability of outcome variables

Source: Author's own elaboration based on fieldwork primary data.

Moreover, farmers from the sample have a high degree of trust in the leadership of their cooperatives (only 18% have none or little trust), probably because they tend to be clustered in the same social communities. Likewise, only about 20% of farmers have a migration background, with the leading reasons for migration being family affairs and perceived better economic opportunities in the destination.

Concerning farm production, most farmers (> 85%) receive their agronomic inputs either for free or subsidized from the GoT, which might explain the finding from the sample that the surveyed AP firms only sign marketing contracts (see Section 3.3) – nevertheless, very few received fertilizers. The sample is almost evenly distributed between farmers who received and did not receive agricultural training; this specific training and all others were vastly offered by either the GoT or the processing firms, or even a combination of both – cooperatives and non-governmental organizations are minor actors. Insufficient training in agricultural best practices affects three-quarters of the SSFs. Regarding potential climate change impact, nearly the entire sample faces serious issues with tree pests and diseases. Moreover, although only 30% face problems with long droughts and excessive temperatures, over two-thirds suffer from unpredictable and changing rainfall patterns. Despite these challenges, two-thirds of the farmers plan to expand their cashew nuts cultivation because of expected high profits, but the lack of available arable land and/or difficult access to new land plots, combined with difficulties in increasing yields, limit these investments. Lastly, almost no SSFs process cashew apples, an activity that would provide them with extra revenue; although some cite concerns about the lack of demand for cashew apples, the vast majority do not know how to process it, which is an opportunity for future policy action.

3.4.2 Regression results

This sub-section reports the regression results and discusses their relevance in light of the extant literature. First, I discuss the main correlates with CF participation in Table 3A.7, whose estimates come from a probit regression with CF participation as the dependent categorical variable; I include a variable controlling for regional differences across districts to minimize biases stemming from the fact that farmers were sampled across different locations. In line with part of the literature, a negative association between landholding and CF participation is found (Bellemare & Lim, 2018; Muriithi & Kabubo-Mariara, 2022; Wang et al., 2014), which suggests that the fact that CF farmers have slightly smaller plots might be driving them to compensate by engaging with CF in the expectation of higher profits. Another potential explanation is that cashew farming is so risky that farmers use only part of their plots of land. Soil quality and soil steepness are not important predictors as SSFs are sampled from similar agroecological areas. No statistically significant correlation exists between own-family labor and adding additional hired workers, possibly because its effect is already captured by farm size and because the decision to enter CF precedes the decision of how much labor input to employ. The same holds for the application of pesticides because farmers increase their application only after entering into contractual agreements due to pre-existing credit constraints (only 13% of the sample has outstanding loans). Farmers with less access to credit are more driven to CF but the schemes currently in place in Tanzania do not provide credit to farmers, which is a possible option for policymakers to explore. Although household size does not affect CF participation, the age of the household head and their gender are negatively associated, indicating that younger farmers and female-headed households tend to be excluded from higher-value market integration, probably because of institutional and cultural reasons. Finally, distance to the nearest tarmac road is positively associated with CF participation[46]. This result seems counter-intuitive – albeit not uncommon as reviewed by Wang et al. (2014) and found by Dubbert (2019) in terms of distance from the farmer's house to the farm –, but it makes sense in the Tanzanian context where all cashew nuts farmers are legally obliged to work with cooperatives and sell via the auction system, thus, access to a buyer is not an issue. Therefore, those farmers farther away from the road networks might be more driven to CF participation to fetch prices higher than those obtained from CBT auctions.

[46] This is not included in the set of controls due to multicollinearity with the district term.

Table 3.3: IV model regressions

	(1)	(2)	(3)	(4)	(5)	(6)
		RCN output			HFIAS	
Contract farming	234.16**	163.81**	172.69**	-0.576***	-0.612***	-0.577***
	(114.15)	(80.86)	(76.62)	(0.118)	(0.153)	(0.096)
Agronomic controls	YES	YES	YES	YES	YES	YES
Household controls	NO	YES	YES	NO	YES	YES
District term	NO	NO	YES	NO	NO	YES
Observations	339	339	339	339	339	339

Note: Standard errors in parentheses: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors are clustered at the village level. Constants are omitted in this table but included in the full models.

Table 3.3 presents regressions for the IV model, which is the preferred specification, adding controls in a stepwise manner. The idea is to test the robustness of the model by adding the relevant controls one block after the other: the results remain stable across the set of controls.

Table 3.4: Summary effects of contract farming on outcome variables

	(1) OLS	(2) PSM-ATT	(3) IPWRA-ATE	(4) IPWRA-ATT	(5) IV
Cashew nuts output	241.22**	318.45*	178.21*	358.74***	172.69**
	(112.45)	(171.75)	(104.88)	(116.33)	(76.62)
Food insecurity (HFIAS)	-0. 637***	-0.705***	-0.692***	-0.770***	-0.577***
	(0.084)	(0.133)	(0.077)	(0.124)	(0.096)
Household controls	YES	YES	YES	YES	YES
Agronomic controls	YES	YES	YES	YES	YES
District term	YES	YES	YES	YES	YES
Observations	339	339	339	339	339

Note: Standard errors clustered at the village level in parentheses: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Cashew nuts output is measured as kg, and HFIAS is an index of 0-4. Constants are omitted in this table but included in the full models.

Next, Table 3.4 summarizes the main results of the effect of CF participation on the outcome variables and it is subdivided into five columns, each with a different estimation strategy – full results are shown in Tables 3A.8 and 3A.9. The preferred specifications are on columns 3 and 6 given that they include the most complete set of controls. They are placed in sequence related to the likelihood of the estimator being more prone to biases: OLS, PSM, IPWRA, and IV – both the ATE and ATT are calculated using the IPWRA estimator to clarify whether a difference exists between the effects considering the entire sample and only the treated (CF) sample. In short, both treatment effects follow the expected signs and are statistically significant in all estimation techniques thus suggesting the result that CF participation increases RCN output and decreases food insecurity is robust. Before discussing the regression results in detail, the following diagnostics are reported:

1. The common support assumption holds as no observations are dropped from it (see Figure 3A.3) ensuring that adequate matches were found and that there is sufficient overlap in the characteristics of treated and untreated units;
2. The PSM estimates are robust to bias stemming from unobservables up to the factor 1.8 (Ichino et al., 2008);
3. The overlap and positivity assumptions hold in the IPW and IV estimations, respectively, as there is no observation with a probability of CF participation beyond the thresholds: 0.001 < $\Pr(CF_i = 1|X_i)$ < 0.999.
4. The overidentification test shows that the H_0 (covariates are balanced between CF and non-CF households) cannot be rejected for all outcome variables.

Regarding RCN output, the two matching algorithms employed produce ATTs with comparable magnitudes. Moreover, notably, the estimations with IV and the ATE are smaller than the OLS baseline, which implies that (1) selection on observables is upward biased, and (2) the ATE is smaller than the ATTs. The latter finding indicates that the effect of CF participation in the treatment group exceeds the effect that participating in such higher-value market linkages would have for the entire sample, which is consistent with Soullier & Moustier (2018) and Ma & Abdulai (2017), but not with Bellemare et al. (2021) and Mishra et al., (2018). This result is expected because of some degree of selection bias that is better addressed with panel data or, preferably, by randomizing treatment assignments, which is not possible in an observational study (Arouna et al., 2021). Analogously, the ATTs of food insecurity are larger than both the benchmark OLS and the ATE, indicating that CF participation tends to benefit more those that select into it compared to the entire sample. In sum, CF participation increases RCN output by 0.13 standard deviations and decreases food insecurity by 0.40 standard deviations in the IV model; this is the preferred specification because it better controls for selection on unobservables.

In addition to the different estimators, I also conduct two robustness checks[47]. First, the ATE and ATT results remain robust when an alternative to the IPWRA model is used, i.e., the augmented-inverse-probability-weighted estimator. Second, in the IV model I include the instruments with different radiuses in a stepwise manner into the regressions: (1) 10 and 100 km; (2) 10 and 200 km; (3) 25 and 100 km; (4) 25 and 200 km. These models are also tested with and without the distance to the nearest CF farmer and all results remain robust across the different specifications.

Finally, four pieces of evidence suggest that the price channel might be a mechanism driving some farmers to participate in CF: (1) a significant majority complain about a lack of reliable buyers and a lack of direct contracts with buyers; (2) the correlation between CF participation and preference to sell RCNs through the auction channel is strongly negative; (3) approximately 90% of the farmers that prefer to sell to AP firms indicate higher prices as a reason; (4) immediate payment and lower transportation costs do not influence those farmers who prefer to sell to AP firms, probably because cargo hauling in the CF schemes of the sample is a task undertaken by the processors or arranged by the cooperatives for the auctions. These factors corroborate the hypothesis that the mechanism whereby CF increases SSFs' welfare is by reducing market risk and partially insuring them against low output prices. The effects are likely a low benchmark because the comparator for CF is the government-run auction system, which is less risky than spot markets, that is the most common comparison used in CF studies (Bellemare & Bloem, 2018).

[47] Results are available upon request.

3.4.3 *Qualitative firms' survey results*

This sub-section analyzes qualitatively the relationships between SSFs and cooperatives and AP firms, with a focus on the latter.

The overall capacity utilization of the firms' sample is quite low at 40% but a bit above the average of the country's processing industry which stands at 35.6%[48]. Generally, capacity utilization falls in periods of economic crisis or low demand, though when it persists low over time it represents a symptom of a chronically inefficient sector that cannot meet the market demand. The main result is the waste of productive resources as a significant portion of the immobilized capital remains idle and still subject to depreciation, on top of underutilization of the accompanying labor factor of production. This situation could be improved with fine-tuning of sectoral policies.

The qualitative survey also shows that most AP firms are investing or planning to invest in new machinery (70%), but only half of these consider more efficient production processes (35%) and a wider variety of final products (25%) in order to functionally upgrade. Regarding capital equipment, the vast majority (85%) employ either manual or semi-mechanized processing technologies and own used rather than new machines, with the only two foreign-owned firms in the sample employing automatic machines with serious plans of acquiring more to increase productivity – the main barrier is access to and cost of finance. Although second-hand equipment can, in principle, be advantageous to some LDCs, mainly because it is cheaper and more suitable to their particular needs, it also represents that cashew processors in the country are far from the technological frontier of the sector – an even worse situation compared to the

[48] Our sample average is close to the national one because we analyze data of all processing firms that are registered with the CBT and, therefore, practically all mid- or large-sized firms. The country figure is from interviews with CBT.

Tanzanian textile industry (Saha et al., 2020).

About a third of the firms in the sample employ a mix that can be considered semi-mechanized, particularly in less labor-intensive activities such as cutting, and peeling. Most of these cases come from firms that are leasing technologically outdated plants from former state-run processors. However, manual labor is still required for calibration, grading and sorting (apart from oversight of machine operations); these activities are not fully mechanized as in the leading country in the sector, Vietnam (Kilama, 2013). Despite recent improvements in the quality and efficiency of machines, Tanzanian firms still experience a lower breakage rate with manual processing; effectively, one manager interviewed stated that their firm invests in machines only if they are supervised by trustworthy staff, otherwise they would experience a large number of broken cashew nuts, thus preferring manual shelling. Consequently, as the most valuable cashews are those that are sold whole, without broken parts, many factories choose to process the higher-grade nuts with semi-mechanized machinery to ensure maximum retention of quality and, thus, profit margins. With such relatively high labor intensity, there are opportunities for improvements of the skills level in the value chain that could result in higher wages and employment levels.

Although the labor factor of production is not an issue for the AP firms, some have difficulties in finding medium- and high-skilled employees for management and complex operational tasks, which hampers mechanization efforts. Although more than half of the firms conduct regular employee training, they are usually very basic, mostly focused on fire and safety regulations and sometimes food hygiene standards. Firms lack the capacity or the financial resources to conduct more complex training, a gap that the GoT could fill with more targeted action from the Vocational Education and Training Authority.

Finally, Tanzanian processors face two main problems in competing internationally. First, expensive and sometimes unavailable credit, blocks the acquisition of new machinery and advance purchases of higher quantities of RCN to increase throughput and, ultimately, capacity utilization (Binswanger & Rosenzweig, 1986). Second, poor infrastructure, a complex tax system, and confusing and/or changing policies and regulations hinder the stable growth of the sector. The few firms that broke through international markets mostly export to Arab countries – a traditional market due to historical ties with Tanzanian traders – and the European Union.

3.4.4 Discussion and limitations

The usual problems documented by the CF literature may be an explanation for the challenges of translating contractual relationships into welfare gains: asymmetric information about farmers' productivity and ability; side-selling incentives for farmers; and contract breach incentives for firms (Barrett et al., 2012; Bellemare & Bloem, 2018; Meemken & Bellemare, 2020). Relational contracting, i.e., repeated interactions between the two contracting parties, is a way to ameliorate these contractual enforcement issues – naturally this is only a long-term solution that requires some level of mutual trust to first initiate and then continue the relationship (Macchiavello & Morjaria, 2015; Michler & Wu, 2020; Ruml & Qaim, 2021b). Macchiavello & Morjaria (2021) found that the legal enforcement of contracts – which is not done by the GoT in the case of CF – is fundamental for the processor-producer relationship to be mutually beneficial, especially in a high competition scenarios. Although the Tanzanian RCN processing market is thin, it is likely to become overcrowded in the case of a poorly planned sectoral expansion, as planned by the GoT, thus harming CF in benefit of poorly defined relational contracts. Moreover, resource-providing contracts – which are absent in our case study – tend to have higher welfare effects than marketing contracts (Dubbert & Abdulai, 2021; Ruml et al., 2022; Ruml & Qaim, 2020), making this a potential policy avenue that the GoT could explore. Further studies using longitudinal data and focusing on resource-providing contracts in different contexts are highly welcomed in this context.

The literature focusing on low- and middle-income countries exploring the relationships between SSFs and large multinational agrifood companies is noteworthy. Farmers are often exploited by draconian contracts and excessive dependency on a particularly large company that provides all services (e.g., financing, inputs, and technical assistance) and acts as the main buyers (Key & Runsten, 1999; Oya, 2012; Wesz Junior, 2022). The survey did not find evidence that this case plays a significant role in the Tanzanian cashew nuts sector because of four factors, some of them similar to the case of the South African sugar industry (Sartorius & Kirsten, 2007): (1) low scale of production, which disincentivizes the capitalization of agricultural activities; (2) considerable influence of cooperatives' leaderships on the contracts signed by farmers; (3) close supervision by the CBT of CF schemes; and (4) despite late payment complaints, which might be caused by some lack of trust, few farmers resent imbalances in their contractual relationships that are so high to the point of blocking an agreement. In fact, farmers complain much more about low prices offered in the CBT auctions, leading to rejections of these price bids and, thus, of repeated delays in selling the produce.

Furthermore, the most common difficulties that SSFs face are related to input and output prices: the former are too expensive and/or not delivered in a timely fashion, whereas the latter are too low and volatile – similar to Kenyan horticulture (Ola & Menapace, 2020a). Although from a different context and crop (maize in Zambia), evidence shows that late delivery of fertilizer reduces yields, which might discourage SSFs from following the best agronomic practices (Namonje-Kapembwa et al., 2017). SSFs often complain about the uncertainty of RCN prices, which is correlated with the preference to participate in CF schemes; evidence in the literature points to CF participation as a means to insure against income variability derived from volatile prices (Bellemare et al., 2021).

Lastly, I discuss the main limitations of this study. First, CF participation cannot be fully randomized, thus primarily collected observational data were utilized with an identification strategy that attempts to tease out the direct effect of CF on welfare outcomes whereby it minimizes selection bias and omitted variable bias. Nevertheless, I remain cautious about the internal validity in terms of claiming causality and, thus, this study's conclusions are limited to reporting associations suggestive of causality. Second, although the sample size is large enough for statistical inferencing, the fact that the sample is split into treatment and control groups reduces the sample size and may impact the efficiency of the point estimates. Third, heterogeneous individual effects within a household could not be identified, given that the survey was at the household-level. And, finally, despite the transferrable characteristics of the sample of Tanzanian cashew nuts farmers to other cash crops growing households in SSA, I refrain from claiming external validity and call for further studies in different countries, sub-regions, and crops, particularly with panel data. Specifically, owing to the structure of the cashew nuts sector in Tanzania, the dataset collected encompasses only family farms and not large corporate farms, which may limit the transferability of the results to regions with large-scale commercial cash crop production.

3.5 Conclusions and policy recommendations

Small-scale farmers in Tanzania who depend on cashew nuts for their livelihoods face high natural risks (pests and weather variability) and market risks (price volatility and guaranteed market for the output) to their production, resulting in low yields. In the other link of the value chain, agro-processing firms need to cope with unpredictable policies, expensive credit, and poor access to international markets. Still, given the importance of this cash crop, the cashew nuts value chain holds the promise of transforming livelihoods in Tanzania. However, it is currently underperforming relative to its potential not only in terms of processing capacity, but also in welfare effects for the rural population. This study presents estimates of the effect of a modern but still incipient market linkage (i.e., contract farming) on SSFs' outcome variables compared to the auctions marketing channel and investigates the interconnections between SSFs and AP firms in the Tanzanian cashew nuts sector. Although I refrain from claiming external validity, the findings may have meaningful implications beyond the study population in Tanzania, particularly due to its structural similarities with other rural African contexts.

The results suggest that policies promoting CF schemes in the Tanzanian cashew nuts sector can reduce transaction costs and contribute to increased output and food insecurity, with estimated standard deviation changes equal to 0.13 and 0.40, respectively. These results can ultimately alleviate rural poverty, which is a pervasive problem in the country. Furthermore, using the qualitative modules of the survey instrument, the following mediation channels for this positive effect are identified in order of importance: a stable market for the raw produce that reduces crop marketing risk; higher prices paid to farmers; consistency of quantities supplied beyond the harvesting season; and higher perceived quality of the contracted RCNs – all captured by the conceptual framework in the context of CF as a means of reducing transaction costs (Dorward, 2001). The effects identified in this study refer only to marketing contracts, as contracts that provide resources, such as inputs and credit, are not observed in the sample.

The methodology used to estimate how CF improves SSFs welfare adopts PSM (Cahyadi & Waibel, 2016) and IPWRA to estimate benchmarks with balanced covariates, controlling for the most important socioeconomic and agronomic factors. Then, instrumental variables improve the identification strategy by addressing self-selection on unobservables and reverse causality. The strength of this chapter’s approach is to contrast results using different methodologies that support the same conclusion. Nevertheless, this study has two limitations. First, the data employed is temporally limited in the sense that the availability of longitudinal data would improve the estimates of the CF effects. Second, the estimates reflect answers from the household head as it was not possible to survey all individual household members owing to practical issues.

The findings suggest that contract farming could also serve as an alternative to the government-controlled auction system in a way similar to how cooperatives can collectively enhance quality standards and mitigate crop marketing risks in a setting plagued with market failures. The current auction model underperforms CF with respect to key household outcome variables (i.e., cashew nuts output and food insecurity) because of its binding association with an inefficient government-controlled auction system as the main marketing channel. Therefore, as a step in the ongoing evolution of food systems (Reardon et al., 2009), vertical coordination between SSFs and AP firms through CF arrangements is important to the upgrading of the Tanzanian cashew nuts value chain. However, much more improvement is needed to further catalyze the process of structural transformation (Bellemare & Lim, 2018).

In light of this chapter’s findings, I propose a set of public policy recommendations focused on actionable measures that the GoT could potentially implement:

1. Increase investments in community-based agricultural extension to provide farmers with the best agronomic techniques to increase yields and incomes;
2. Focus agricultural subsidies on the replacement of old trees and the supply of high-yielding varieties of cashew seedlings – the adoption of high-yielding varieties of pigeon pea in Tanzania, for instance, showed promising results (Asfaw et al., 2012);
3. Ease access to land leaseholds and to the options that the farmers’ lessees have with their titles to drive new investments into RCN production;
4. Increase investments in vocational and educational training to ensure that the workforce of processing firms has adequate qualifications for value chain upgrading;
5. Support small-scale processors with subsidies targeted to strict sunset clauses to integrate into well-functioning supplier–buyer networks;
6. Conduct studies about the financial feasibility of subsidizing credit and providing loan

guarantees through parastatals to facilitate local processing firms to purchase RCN so that they can increase capacity utilization throughout the year.

CHAPTER 4

The effect of trade and customs digitalization on agrifood trade: A gravity approach[49]

4.1 Introduction

Agrifood trade offers great potential for LMICs to create additional income and to diversify their export structures into non-traditional export sectors, particularly processed agricultural products, which are either largely produced in high- and middle-income countries or those markets are the destination of the meager LMICs exports of processed agrifood (Klasen et al., 2021; Kornher & von Braun, 2020). Largely based on consumption of locally grown crops, with the exception of some grains, the food systems of LMICs undergo a structural transformation towards: (1) higher capital intensity, (2) larger scale of production, (3) deeper integration to GVCs, and (4) lower dependence on subsistence agriculture in benefit of crop commercialization (Barrett et al., 2022). Despite the potential from natural endowments and comparative advantages in some agrifood products, agrifood trade in SSA is still comparatively low.

The launch of the African Continental Free Trade Area (AfCFTA) at the beginning of 2021 increased the political momentum for deepening regional trade agreements (RTAs) – in particular harmonization of customs procedures and reduction of tariffs and NTMs among regional trade partners – and also bears great potential to redefine and improve SSA's role in global agricultural trade (Fofack et al., 2021). Notwithstanding recent integration efforts, intra-African agrifood trade still represents a small share of continental exports (~20%) and imports

[49] **Publication status:** Bueno Rezende de Castro, A., Kornher, L. (2022): Published on *Q Open* <https://doi.org/10.1093/qopen/qoac037>

(~15%), with strong regional differences and, in particular, with a lower share of processed products vis-à-vis unprocessed (Bouët et al., 2021). Consequently, despite a gradual deepening of intra-African trade and a more prominent role played by higher value domestic markets, in the medium-term there is more value to be accrued by tapping into international food markets, especially from HICs. The two reasons for this are that HICs have a higher demand for processed products and these products have higher income elasticity compared to unprocessed ones (Fukase & Martin, 2020). Moreover, international trade has the potential to alleviate hunger situations via matching food surplus regions with food-deficit regions as well as contributing to climate change adaptation (Janssens et al., 2020). Hence, international trade is an important pathway for SSFs in LMICs to improve livelihoods by increasing food commercialization, constituting an important developmental step for these countries to integrate their agricultural production to trade flows as a pre-condition to increase value addition (Barrett et al., 2022).

Transaction costs, spanning from poor physical infrastructure to deficient e-payment systems, represent a key determinant of international trade performance. Trade costs are substantially higher if developing countries are involved and up to five times higher between SSA or North African countries than between European Union (EU) or North American countries (United Nations, 2019); there is also evidence from intranational data that agricultural trade costs within SSA are up to five times higher compared to other regions in the world (Porteous, 2019). Hence, reducing transaction costs could significantly increase agrifood exports from SSA. This chapter focuses on another type of less explored transaction costs, which are quickly becoming a major impediment to international trade: logistic and bureaucratic costs related to transportation and customs procedures of goods and services (Zaki, 2010, 2015). In addition to a diverse array of actors and different and complex technical requirements, one of the most stringent barriers to international trade are customs bureaucracies, which, in general, are country- or region-specific, thus susceptible to a plethora of standards – in particular more demanding in South-North trade

(Fiankor et al., 2021; Sun et al., 2021). These formalities occur in four stages of procedures: commercial, transportation, regulatory, and financial (Civelek et al., 2017). In this sense, they represent a key factor in the positioning and competitiveness of countries in regional value chains (RVCs) and GVCs (Ma & Van Assche, 2011). Tariff reduction was central to all WTO negotiation rounds and many least developed countries (LDCs) in SSA are granted free access to HICs through preferential trade agreements or initiatives with unilateral trade preferences[50]. The challenge is then how to take advantage of subsequent waves of trade liberalization, less reliant on tariffs reduction, to not only expand trade volumes but mainly increase value addition.

This chapter investigates empirically the hypothesis that trade digitalization measures increase bilateral agrifood trade, particularly for more complex products with higher value addition, and especially for SSA and Asian exporters – the causation chain ends with an increase in economy-wide productivity, which is the main determinant of long-run growth. The digitalization processes considered here look at trade facilitation through a reduction in transaction costs related to logistics and bureaucracy, that can be implemented more widely, in terms of general procedures and processes, or more specifically, in terms of facilitating compliance with NTMs. Trade and customs digitization is defined as the substitution of physical documentation of regular trade procedures (e.g., payments, auditing, the release of cargoes, certificates of compliance, etc.) by the application of modern ICTs. These include, (a) paperless and electronic digitized tracking of the trade procedures (Duval et al., 2019), henceforth referred to as PT, and (b) the digitalization of cross-border customs and financial procedures of NTMs, or CB. Both can facilitate and expand international trade flows, especially in the current context of a retreat of globalization by reducing the bureaucracy costs in the relationship between public authorities and citizens and companies (Reil et al., 2022).

[50] The main examples are the European Union's "Everything but Arms" or the North American African Growth and Opportunity Act.

In light of this situation, that increases transaction costs in international trade, WTO members adopted in 2017 the TFA, which seeks to cut red tape at the two border points via modernization, simplification, and harmonization of customs procedures. Currently, the implementation of the TFA stands at only 41.3% for LDCs (the majority of them in SSA), while for developing countries in general it is almost double[51]. What is more, the WTO TFA mostly crystallized improvements already underway rather than pushed the trade facilitation agenda further. The ultimate objective is to deliver the promise of the digitalization of trade procedures and speed up its unrolling in the developing world; in this regard, however, SSA is threatened to be left overtaken by trade competitors due to a slower pace of adoption (Ibrahim et al., 2019). According to Venables (2004), if the trade costs of a small open economy (approximately what Tanzania is) falls relative to the rest of the world, this country would become more internationally competitive in more products; but which type of products gain more would determine if its comparative advantages would shift or not.

Although trade and customs digitalization are not a silver bullet to shift comparative advantages and boost an export-led growth strategy, paperless mechanisms and procedures can facilitate and expand international trade flows, especially in the current context of a retreat of globalization. This study adopts two main variables of interest: PT covers the digitalization of transportation, regulatory, and customs procedures in general; CB deals with the digitalization of customs and financial procedures of NTMs. Regarding the introduction of PT measures, SSA lags behind other competing markets and target destinations, such as Asia and the EU (Laryea, 2005), thus there is ample potential for improvement.

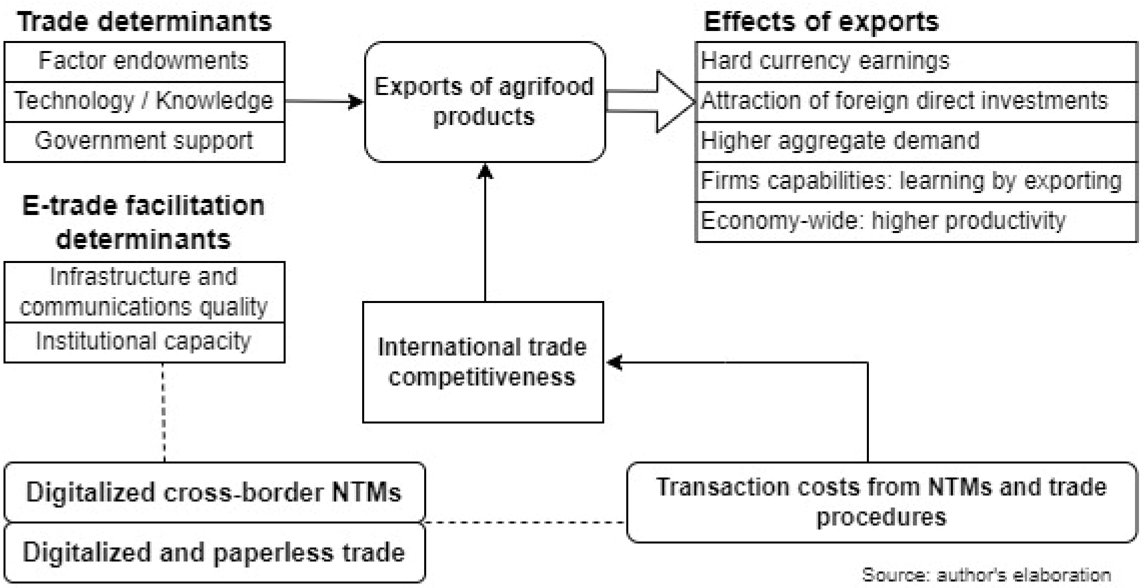
This chapter contributes to the literature in three ways: (1) based on Transaction Costs Economics (Williamson, 1979), I elaborate how trade digitalization can reduce transaction costs and, thus, affect international competitiveness in agrifood trade; (2) by combining two rich

[51] https://tfadatabase.org/ <Accessed on 10 November 2023>.

datasets[52] spanning three years (2015, 2017, and 2019, but not the respective middle years), to the best of my knowledge, this is the first study to employ a gravity model approach to estimate the effects of e-facilitation of trade procedures and bureaucracies on agrifood trade; (3) I provide granular results with respect to different agrifood products and the sub-regions, thus filling the literature gap on how trade and customs digitalization can benefit specific agrifood sectors and regions of the Global South. The methodology employs a gravity model, which is the workhorse empirical methodology to estimate the effect of policy variables on bilateral trade flows.

The remainder of this chapter is organized as follows. Section 4.2 follows this Introduction with the background of the problem by framing a literature review into the conceptual framework. Then Section 4.3 describes the data and explains how the empirical methodology addresses endogeneities. Section 4.4 presents and discusses the results and Section 4.5 concludes the chapter with several relevant policy implications and some limitations of the analysis.

4.2 Background and conceptual framework



[52] See Section 4.1 for more details.

Figure 4.1: Conceptual framework: cutting transaction costs related to trade and customs digitalization increase agrifood

This section has three objectives: (1) introduce the conceptual framework this chapter employs, which explains the causal pathways of trade digitalization measures on bilateral agrifood trade flows (see Figure 4.1); (2) place the literature review in the context of the conceptual framework; (3) discuss the microeconomic foundations of the gravity model.

This chapter is embedded in the overall framework of this dissertation following transaction cost economics, which does not assume that supply and demand always clear without friction, due to the presence of transaction costs (Williamson, 1979). Given the need for LMICs to catch up with production technologies, i.e., the production technological frontier, the competitiveness of a firm or a country in international trade is largely due not only to productive efficiency but also due to transaction costs. This chapter adopts two definitions: first, trade facilitation is understood as the simplification and harmonization of trade processes related to sending, receiving, and processing data, documents, and other information required for international trade (Engman, 2009; Grainger, 2011); second, directly related to the previous concept, e-trade facilitation simply makes all of these processes and information related to international trade electronically, and often remotely, available for customs, other official authorities, and the trading parties, thus reducing transaction costs (Duval & Mengjing, 2017; Lewis, 2009). Naturally, as the bedrock of all digitalization efforts, the introduction of the internet increases trade via lower information costs (Lin, 2015).

When trade facilitation is hampered by bureaucratic hurdles as a result of inappropriate and/or inefficient digitalization, firms face longer waiting times to clear goods and process all paperwork (either at the origin and/or at the destination), which reduces firms' and countries' competitiveness in the GVCs via the channel of increased transaction costs. Given that there is a lack of specific data that prevents a direct measurement of transaction costs and, thus, of its impact on trade, PT and CB are assumed to cover the majority of the transaction costs associated with customs and trade procedures, which is the research object of this chapter. Hence, by estimating the effect of these digitalization procedures on bilateral agrifood trade, this chapter provides an approximation of the burden imposed by transaction costs on international trade.

Figure 4.1 summarizes the relationships between the trade digitalization measures captured in this study and trade competitiveness. A group of factors related to the diamond model (Porter, 1985) determines the performance and competitiveness of countries in international export markets, mostly related to the production process (i.e., technology and factor endowments) and the enabling environment (i.e., ICTs, infrastructure, institutions, and government support). Higher exports are positively related to overall economic growth (Frankel & Romer, 1999; Gözgör & Can, 2017) – especially so agricultural exports in LMICs (Dawson, 2005) – despite evidence challenging the export-led growth model for SSA (Were, 2015), therefore I expect that the e-trade facilitation role of the policy variables leads to better developmental outcomes. The mediating channel between trade facilitation via digitalization and higher agrifood trade are transaction costs, specifically those related to trade and customs digitalization measures, represented by the two key policy variables, PT and CB. Although this is outside of the scope of this chapter, it is also likely that the reduction of transaction costs makes international prices clear supply and demand more efficiently, thus reducing distortions in agrifood markets.

In the agrifood sector, NTMs and associated compliance bureaucracy are one of the main contributors to transaction costs. NTMs involve both technical and non-technical barriers to trade, e.g., sanitary and phytosanitary standards (SPS) (Melo & Shepherd, 2018). Restrictive food standards are often justified by reasons related to consumer protection and standardization of product quality, which are associated with increased trade flows depending on the trade partners and products involved (Bratt, 2017; Santeramo & Lamonaca, 2019). Nevertheless, they also might be imposed to protect domestic producers without legal recourse to imposing tariffs (Kareem et al., 2018). Despite the positive demand-driven trade effects of improving product quality and standardization (Santeramo & Lamonaca, 2019; Xiong & Beghin, 2014), NTMs invariably increase the import parity price of agrifood products not only because of their direct effect but also indirectly by imposing transaction costs related to the compliance with and monitoring of these measures (Fiankor et al. 2021; Melo & Shepherd 2018).

Santeramo and Lamonaca (2019) provide an excellent review of the ambiguous literature of the trade effects of NTMs. Xiong and Beghin (2014) argue that, given the lower level of technical capabilities and lower quality of ICTs that LMICs generally have vis-à-vis HICs, transaction costs arising from NTM compliance, coupled with laggard e-trade facilitation, tend to be higher in LMICs and lead to trade diversion. The consequence is that exporting LMICs face decreasing competitiveness in products in which most of them have revealed comparative advantages, such as agrifood products, therefore reducing their exports (Ehrich & Mangelsdorf, 2018).

The three hypotheses that this chapter tests are based on the conceptual framework:

1. PT: Facilitation of e-trade bureaucracy via standardized electronic documentation and processes[53] cuts transaction costs and, thus, has a positive effect on bilateral flows of agrifood trade.
2. CB: Facilitation of electronic processes and paperwork related to cross-border NTMs also cuts transaction costs and, thus, has a positive effect on agrifood trade.
3. Facilitation of e-trade bureaucracy and NTMs has a greater impact when done in LMICs as compared to HICs, because the former are lagging the latter in this aspect, due to their lower income level.

The importance of e-trade facilitation lies in the simplification of shipment, payment, standards verification, and procedures pertaining to customs inspections and clearances on both origin and destination, etc. This bureaucracy related to international trade is of the last-mile type since it is not related to the production process per se, but rather to the last economic activity – selling

[53] Interchangeably referred as “paperless trade” throughout this paper.

and transporting the goods between countries –, and this process can often be especially troublesome due to language and cultural barriers and to first-time exporters and small and medium firms (Nooteboom, 1993). Particularly in LMICs, few firms export large quantities (World Bank, 2020a); in addition to poor business planning and limited access to high quality inputs, which result in lower competitiveness, the lack of PT measures can impose an extra burden, with some firms even resorting to hiring intermediate broker agents. Therefore, one of the easiest ways of facilitating trade flows is to introduce standardized electronic documents and information exchange. The main impact pathway is directly related to the increase/reduction of transaction costs stemming from trade digitalization measures, which then decreases/increases total flows of exports of agricultural products (Goedhuys & Sleuwaegen, 2016; Milner et al., 2000).

For instance, when an LMIC exports a certain agricultural product to a HIC, it often faces high transaction costs related to different and complex procedures to comply with NTMs; standardizing and streamlining these steps can increase efficiency by reducing the trade-related lead time and financial costs (Czubala et al., 2009), although there is evidence pointing towards a negative trade effect of SPS-related NTMs in the agricultural sector (Li & Beghin, 2012). Furthermore, the rapid worldwide spread of the internet in the early 2000s greatly increased the growth of goods' exports (Freund & Weinhold, 2004), which served as the springboard for the recent wave of e-trade facilitation. The dataset employed captures these trade facilitation variables[54] from both the importer and exporter sides, thus the specific impacts of PT and CB on agrifood trade can be more precisely analyzed compared to when measures are taken by only one side of the bilateral trade relationship.

The importance of e-trade facilitation lies in the simplification of the manifold processes in international trade, the main ones related to shipment and payment, in addition to a variety of

[54] The definitions of the policy variables are in the Annex Table 4A.2.

procedures pertaining to customs inspections and clearances on both origin and destination, taxes and tariffs invoicing, technical and sanitary inspections, etc. This bureaucracy related to international trade is of the last-mile type since it is not related to the production process per se, but rather to the last economic activity: selling and transporting the goods between countries. This process can often be troublesome due to language and cultural barriers, the different standards, and requirements, and the additional costs it contains, especially to first-time exporters and small and medium firms (Nooteboom, 1993). Particularly in LMICs, few firms export large quantities (World Bank, 2020a); in addition to poor business planning and limited access to high quality inputs, which result in lower competitiveness, the lack of PT measures can impose an extra burden. Some firms even resort to hiring intermediate agents who assist them with the paperwork and bureaucratic processes (i.e., logistic, legal, administrative, exchange rates, etc.), which naturally increases costs. Therefore, arguably, one of the easiest ways of facilitating trade flows is to introduce standardized electronic documents and information exchange.

There are plentiful studies analyzing the Asia-Pacific region in terms of the impact of trade facilitation measures, which is mainly due to its export-led growth model and trade opening in the 1970s and 1980s, resulting in a proliferation of RTAs (Duval & Mengjing, 2017; Shepherd & Duval, 2016). Kumar Roy & Xiaoling (2020) found evidence that trade facilitation policies introducing PT measures in South and Central Asia increased export performance. Using the same dataset that this chapter employs (see the next section for more details), Duval et al. (2018) and Duval et al. (2019) also found evidence in the same direction for these regions, but focusing on the potential for trade cost reduction of PT measures. South Korea, in particular, was already pioneering the introduction of ICTs into trade formalities in the early 2000s (Yang, 2009). Apart from trade flows, e-trade facilitation was found to also stimulate foreign direct investments (FDI) (Yasui & Engman, 2009). The main pathway for this effect is similar to the one described in this dissertation’s conceptual framework: more efficient procedures for risk management, pre-arrival processing, and post-clearance audits reduce transaction costs. Engman (2009) also found that tax revenues are more effectively collected in the presence of e-trade facilitation measures.

The literature is clear that imposing NTMs, especially SPSs, has negative impacts on trade flows (Melo & Shepherd, 2018; Otsuki et al., 2001). However, many of these SPSs are justified on the basis of animal and human health and will remain in place, thus harmonizing and sharing common SPSs can stimulate trade because the barriers to trade that NTMs impose are outweighed by a reduction in the burden of information asymmetry and transaction costs that such harmonization brings (Moenius, 2004). These explain the role of CB measures to ease the flow of trade. Furthermore, there is an increasing usage of digital technologies to facilitate SPS in agrifood trade, for instance, traceability and supply chain integrity checks, screening for pesticides, advancement consignment declaration, and e-certificates for plants and animal products reinforces (Avery et al., 2021).

With the increasing emergence of GVCs, some theories of trade portray that the exchange of goods and services is majorly driven by countries' need to have access to different technologies, however, the bulk of global trade is concentrated in countries with similar factor endowments and productivity levels (Feenstra, 2015; Fidrmuc, 2004). Moreover, a significant share of trade is done in intermediate products, including agricultural raw materials that are lightly processed in the origin for further processing in the destination (World Bank, 2020a). Although trade in goods has become less sensitive to the distance between countries, given the technological advances that slashed transportation costs, geographical distance is still a strong predictor of bilateral trade flows – as our gravity results demonstrate – and, thus, can increase transport costs and thwart trade (Brun et al., 2005). The negative correlation between distance and trade is particularly strong when at least one of the two trading countries is a LMIC, with the bulk of the positive effects of technological progress and lower transaction/transportation costs being accrued by HICs (Arvis et al., 2016).

In this regard, I turn to the gravity model and place it in the context of the trade literature, where it first appeared relatively disconnected from theoretical groundings (Tinbergen, 1962), however, it has seen significant improvements (Anderson, 1979; Bergstrand, 1985; Philippidis et al., 2013). The main contribution to enhance the gravity model theoretical micro-foundations came from Anderson & van Wincoop (2003) who explicitly accounted for the role of trade costs by incorporating the multilateral resistance term, which are price indices dependent on trade barriers. These can be either natural (transportation costs and time, comparative advantages, etc.) or artificial (usually tariffs and NTMs), and are important determinants of bilateral trade, together with external trade barriers with third parties, which are captured in the model by fixed effect dummies for each country (Anderson & van Wincoop, 2004).

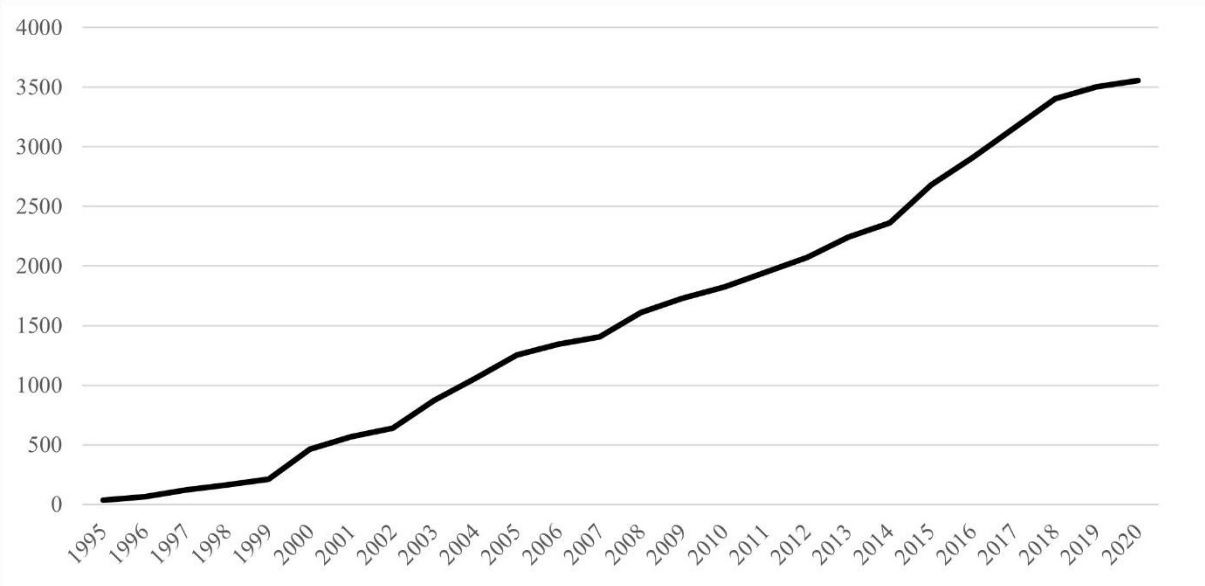
The recent literature has been producing a growing body of evidence on a wide array of transaction costs and other impediments to trade in food and agricultural products; recent developments have also applied the gravity model to this topic (Grant & Lambert, 2008; Lambert & McKoy, 2009; Mujahid & Kalkuhl, 2016; Philippidis et al., 2013; Sarker & Jayasinghe, 2007; Sun & Reed, 2010). Nevertheless, there are still limited studies on the effects of the digitalization of bureaucratic procedures and customs formalities on agrifood trade flows between LMICs and HICs – apart from one study of digitalization of SPSs but considering total trade flows (Avery et al., 2021). This chapter fills this specific literature gap by providing evidence with heterogenous regional effects.

4.3 Materials and methods

4.3.1 Data

The main dataset utilized in this chapter is the International Trade Database at the Product-Level (BACI) from the Center for Prospective Studies and International Information (CEPII) (Gaulier & Zignago, 2010). It provides importer and exporter data with worldwide breadth, then merged with the level of product disaggregation at the 3-digit level of the Standard international trade classification (SITC) nomenclature based on UN Comtrade data – reviewed for its accuracy and consistency. In addition to that, I use the CEPII's gravity modeling data set that encompasses the most relevant gravity modelling control variables. The main trade policy variables (PT and CB) are extracted from the United Nations Global Survey on Digital and Sustainable Trade Facilitation (UNTFS) (United Nations, 2019), which is merged with the BACI dataset, thus yielding a time span of five years, and three points in time: 2015, 2017, 2019. The UNTFS data set includes 125 countries. The full list of traded products this chapter considers broken down by degree of processing is on Table 4A.1.

This chapter conducts two types of analysis: first, an analysis of digital trade facilitation available in the exporting country (country of origin); and second, an analysis of the same digital trade facilitation available in the importing country (country of destination). Naturally, the selection of countries in the first analysis is more restrictive and will not be based on a representative data set for LMICs. In the second analysis, I make use of the fact that agrifood imports are concentrated among fewer (mostly industrialized) countries and, therefore, this country sample is representative. The selection of this set of 128 countries (25 of those in SSA) is motivated by computing the import share they collectively are responsible for: 87%[55].



Source: Author's elaboration based on data from UNCTAD Trains

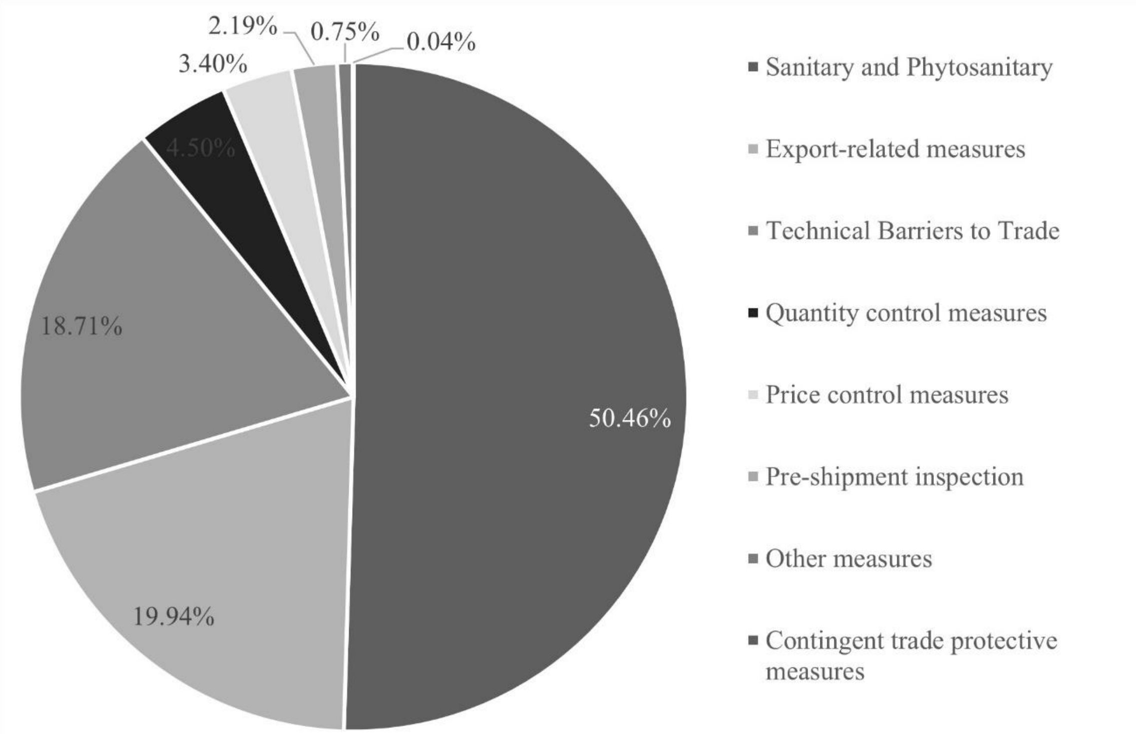
Figure 4.2: Number of total NTMs valid in each year in the agricultural sector – period 1995-2020

The full set of policy variables investigated in this chapter is detailed in Table 4A.2 (United Nations, 2019). They are all involved with the introduction and adoption of ICT-related mechanisms and procedures to simplify trade: PT, CB, transparency, formalities, and institutional arrangement and cooperation. Individual measures have values equal to 0: not implemented, 1: pilot stage of implementation, 2: partial implementation, 3: full implementation. Then, the indices are compiled by averaging those individual measures and their sum is normalized from 0-1. From this list, I emphasize two main variables of interest from the digital trade facilitation sub-group: PT and CB, the latter of which is defined as measures related to the digital exchange and facilitation of specific NTMs-related information and data. Figure 4.2 captures, for the past few decades, the growing number of NTMs in agriculture, the sector most affected by NTMs (Melo & Shepherd, 2018; Niu et al., 2018). From this sector's perspective, digitizing documents related to SPSs and rules of origin certification are prominent and provide the potential to facilitate exports of LMICs. The difference of CB

[55] Data from UN Comtrade.

and PT is that the latter refers to general e-facilitation of trade procedures not related to NTMs, in other words, it encompasses the application of ICTs to trade-related services (United Nations, 2019).

Food standards of major agrifood importers have been identified as a major impediment to agricultural trade (Fiankor et al., 2021; World Bank, 2008). In fact, the definition of standards beyond the international regulation level could be considered a quasi-protection of the domestic agricultural sector in the European Union (Shepherd & Wilson, 2013). Figure 4.3 reports that over half of the NTMs in the agrifood sector are related to SPS measures, and another fifth of NTMs are other export-related measures; both types of NTMs are captured by the CB policy variable.



Source: Author's elaboration based on data from UNCTAD Trains

Figure 4.3: Percentage of non-tariff measures in the agricultural sector by type (10/2021)

4.3.2 Empirical methodology

The identification strategy employs the gravity model to estimate the impact of the policy variables on bilateral agrifood trade. One of the most important problems with naïve gravity modeling is that it does not consider the heterogeneity of the socioeconomic structures of trading countries, thus in order to avoid the omitted variable bias in the estimation, I control for these characteristics; in our case, specific to country pairs since I am using bilateral trade data. The most common of those characteristics are cultural, political, and historical factors; even when controlling for the size of the economy and distance between trading countries, such factors are important co-determinants of bilateral trade flows (Yotov et al., 2016). Namely, the baseline gravity model is expanded by including the following country-specific socioeconomic and cultural factors: contiguity (Bergstrand, 1985); common official language and/or a former colonial relationship, including common colonizer (Frankel et al., 1995); common origin of the legal system; and coverage and type of regional and preferential trade areas/agreements (Grant & Lambert, 2008; Lambert & McKoy, 2009; Sarker & Jayasinghe, 2007; Sun & Reed, 2010). These variables are captured with vector C in Equation 4.1 (original gravity model). The full list of the social-cultural aspects is in Table 4A.3.

$$\ln X_{ijts} = \alpha_0 + \beta_1 \ln GDP_{it} + \beta_2 \ln GDP_{jt} + \beta_3 \ln dist_{ij} + \beta_4 \ln Policy_{ijt} + \beta_5 C_{ijts} + \gamma_i + \gamma_j + t + \varepsilon_{ijts}. \tag{4.1}$$

Here, X is the bilateral agrifood trade flow between exporter i (origin country) and importer j (destination country) at the year t, s is the product SITC code, α is the positive gravitational constant, GDP measures the size of the economies in each country, *dist* is the geographical distance in Kms between the pair of trading partners (a proxy for transport costs), and ε is the idiosyncratic error term. β_4, which estimates the impact of the policy variables, is the key parameter: e-trade facilitation via paperless trade and cross-border NTMs in the pair of trading countries. The decision to model the trade flows at the SITC 3-digit code, instead of product level, is to minimize the share of zero trade observations and to keep the data set manageable without making arbitrary decisions about the products considered in the analysis.

The econometric specification does not directly include measures of NTMs at the destination because they are usually product-specific and difficult to aggregate at the SITC level. Moreover, using ad valorem equivalents may make changes in the policy variable indistinguishable from changes in the level of NTMs. However, the econometric specification indirectly controls for NTMs. That is because the distance between the largest city of each country is included, thus the distance between pairs of countries is a unique value – effectively like a country-pair fixed effect – that captures not only the distance but also the specific trade relations and trade costs between the pair of trading partners. And also, the same models with sub-sets at the product category level are estimated and add other institutional variables, e.g., formalities and transparency, in some specification – both capture elements of NTMs.

Despite the appropriateness of the gravity model, this identification strategy is subject to potential bias from endogeneity. First, there are problems with the lack of a time dimension to allow explanatory variables to adjust (Yotov et al., 2016). In this regard, I make use of the panel characteristic of the data set – term t in Equation 4.1 – in order to consider the time dimension. Second, there is omitted variable bias because the model does not account for all potentially relevant variables either because they are difficult to observe, and/or to quantify. Although this is addressed by adding various socioeconomic controls, some relevant variables might still be missing. One example is NTMs: it is difficult to find a variable for this considering the aggregate nature of the outcome variable (Melo & Nicita, 2018). For instance, unobserved country-specific relevant variables might be correlated with the error term and, thus, impact agrifood trade. To tackle this issue, the baseline gravity model is expanded with the inclusion of the multilateral resistance term, accounting both for time- and country-fixed effects, captured by the terms t and γ, respectively, for the importing and exporting country. The mechanics are simple: I incorporate fixed effects for the opposite country vis-à-vis the country of the policy variable (Cheng & Wall, 2005). Effectively, I swap between destination and origin, so if the policy variable is measured at country i, the fixed effect incorporated into the regression is from country j, and vice versa. This strategy avoids incorporating all country- and time-pairs fixed effects (FE) that absorb most of the variation of the policy variables, thus jeopardizing its identification, since identification comes from variation over time (Cipollina et al., 2016). The time trend also absorbs all macroeconomic shocks common to all country pairs that occurred in the period of analysis. Standard errors are clustered at the distance between trading country pairs to mitigate skewness stemming from missing unobserved variables that may differ systematically by country.

On the other hand, I cannot fully rule out reverse causality. However, similar to the gravity literature (Flach & Unger, 2022; Kabir et al., 2017), I argue that bilateral trade flows do not affect the policy decision of an importing country towards all its trading partners. In particular, this chapter is interested in how digital trade facilitation can improve agrifood exports from developing countries to developed countries. I argue that the policy decision regarding trade digitalization measures especially in HICs, but not exclusively, is not related to the bilateral trade flow from a single developing country.

Despite this extended version, there is still an underlying problem with gravity models in the form of a large number of zero-valued bilateral trade flows that inevitably happen when dealing with one-sided large data sets. This is expected as neither all countries produce all goods, nor all countries trade with all other countries. For instance, Haveman & Hummels (2004) report that 10% of all countries concentrate 58% of trade flows, providing evidence for the hypothesis that zero-valued trade is a normal occurrence and needs to be controlled for, especially in the case of agrifood data (Haq et al., 2013). Merely excluding the zero values is not adequate because the presence of zero values, which are often non-randomly distributed, convey important information about the nature of the trade flows (Burger et al., 2009; Eichengreen & Irwin, 1998). Moreover, excluding the zero values preclude analysis of bilateral trade creation as the sample would only have observations of actually traded goods, and no potential trade – the estimates would be conditioned on actual trade taking place. In the case of this chapter's sample, 83% of the agrifood trade observations equal zero, which is common in the empirical trade literature (Westerlund & Wilhelmsson, 2011), and results in enough positive trade flow observations in all model specifications. Since a log-log model is estimated, I simply add a constant of 1 to zero-trade flows in level in order to minimize any bias.

Furthermore, we adopt a robust solution to deal with the problem of zero-valued trade flows by employing a model that originally uses count data, but adapted to non-negative continuous variables accounting for fixed effects (Anderson & van Wincoop, 2003) – thus, the preferred specification is the PPML expanded with country fixed effects (Silva & Tenreyro, 2006), which solves the problem of the omission of the multilateral resistance terms (Fally, 2015). Since this estimator belongs to the class of pseudo maximum likelihood, the PPML does not make any assumptions about the distributional form of the variables; it simply requires that the conditional mean of the dependent variable is correctly specified (Santos Silva & Tenreyro, 2010; Yotov et al., 2016). This specification has six main advantages: (1) consistent, unbiased, and efficient in presence of heteroskedasticity – as it does not assume that the error variance is constant across observations; (2) consistent in the presence of over- or under-dispersion as it makes no assumptions on the dependent variable distribution; (3) appropriately deals with zero-valued trade flows due to its multiplicative form, resulting in a positive mean; (4) all observations are weighted equally; (5) better avoids sample selection bias; (6) produces estimates in which actual and estimated trade flows are equal across all bilateral trade partners (Arvis & Shepherd, 2013; Santos Silva & Tenreyro, 2010, 2011; Yotov et al., 2016).

The main difference to the other methods of gravity model estimation, such as the negative binomial, is that the zero-inflated PPML model accounts for two states with probability p_i: excess of zeros, indicating no trade of a particular product between a pair of countries, or the probability of trade, $1 - p_i$. This approach considers the true zero values processes that capture the non-existence of trade, but with a non-zero probability of trade, as opposed to the false zero processes that simply record all zero trade values. In short, this means that a pair of countries might not be trading a particular product due to demand and supply conditions as well as other precluding factors, but this does not mean that trade cannot happen in this context. Consequently, the class of PPML models performs well in presence of large numbers of zero (Santos Silva & Tenreyro, 2011). Given this rationale, the PPML estimator is employed in Equation 4.1 and I drop a baseline estimated with OLS because it is not a consistent estimator in our case.

4.4 Results

4.4.1 Descriptive statistics

This sub-section starts by providing an overview of the policy variables. Figures 4.4-4.6 capture the adopting countries' relative frequency of the degree of implementation of three policy variables across the time span of the panel: total measures (measured from 0-1.0), PT (measured from 0-0.3), and CB (measured from 0-0.2), respectively. All charts are sub-divided by year of the panel and by origin/exporter vis-à-vis destination/importer. I highlight three patterns from these figures: (1) The distributions of CB are skewed to the left, which means many countries have not yet adopted such measures, and those that adopted have not fully implemented them, considering the low scores; (2) PT and the variable capturing all measures combined (total) have a higher degree of implementation compared to CB; (3) The graphs show some improvements as the curves become less skewed to the left – i.e., countries started to implement the measures –, which means that, despite some recent progress, there is ample scope for improvement in e-trade facilitation, especially regarding NTMs; (4) Another potential explanation for the lack of implementation of the e-trade facilitation measures is that the majority of the adopted sample comprises developing countries, which are in early stages of e-digitalization, while developed countries face difficulties and high costs, such as sunk costs, in adapting their old analogical systems (United Nations, 2019).

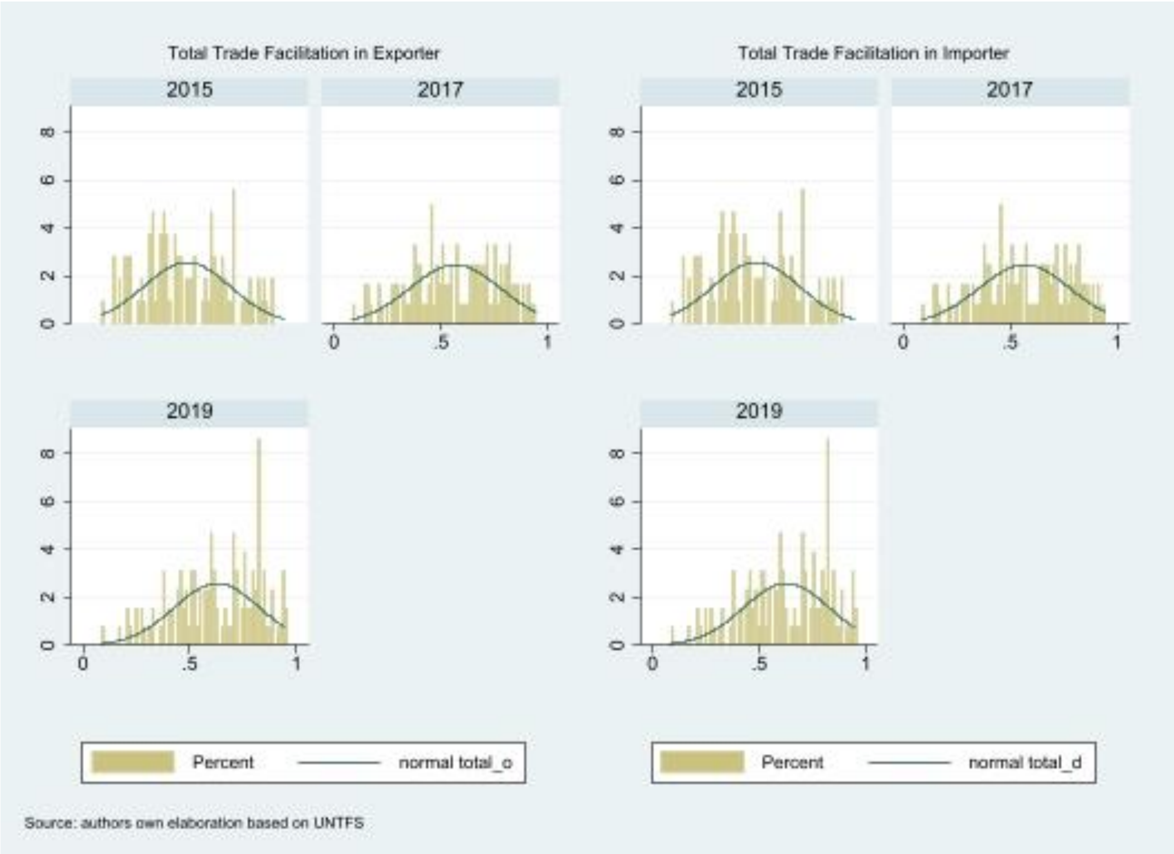


Figure 4.4: Distribution of country-level total trade facilitation index

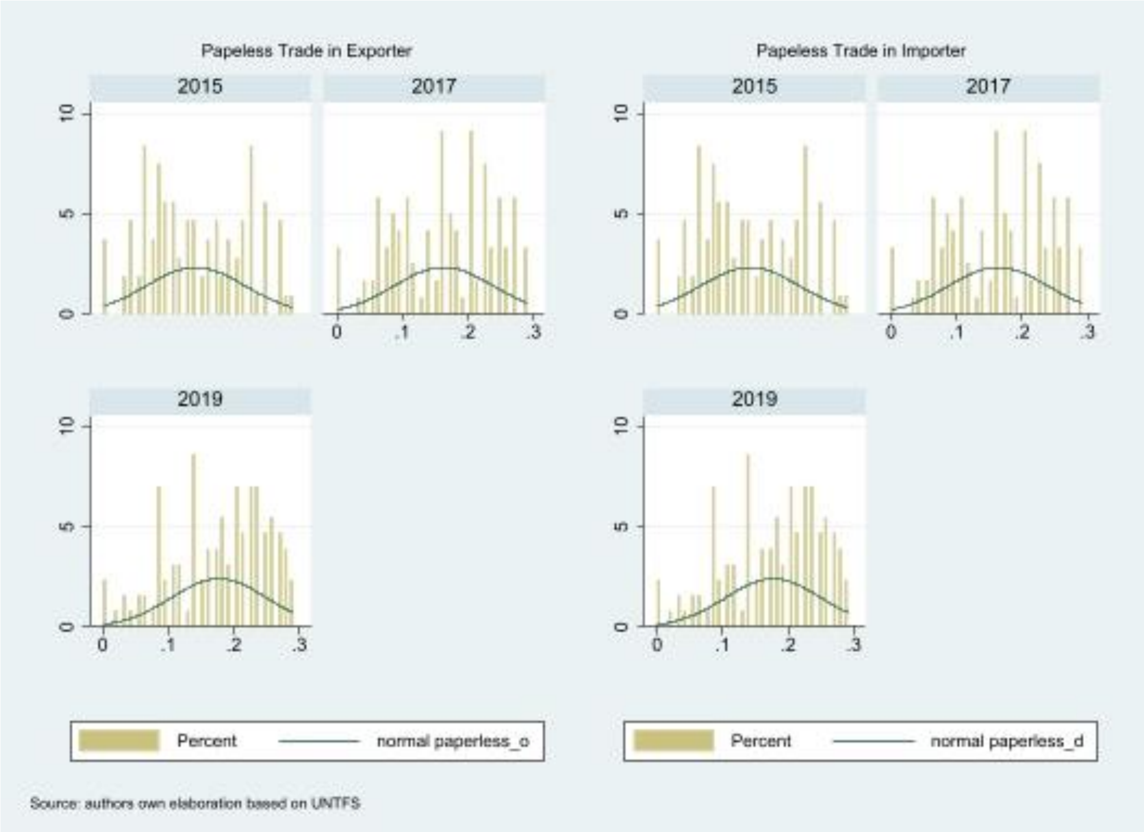


Figure 4.5: Distribution of country-level paperless trade index

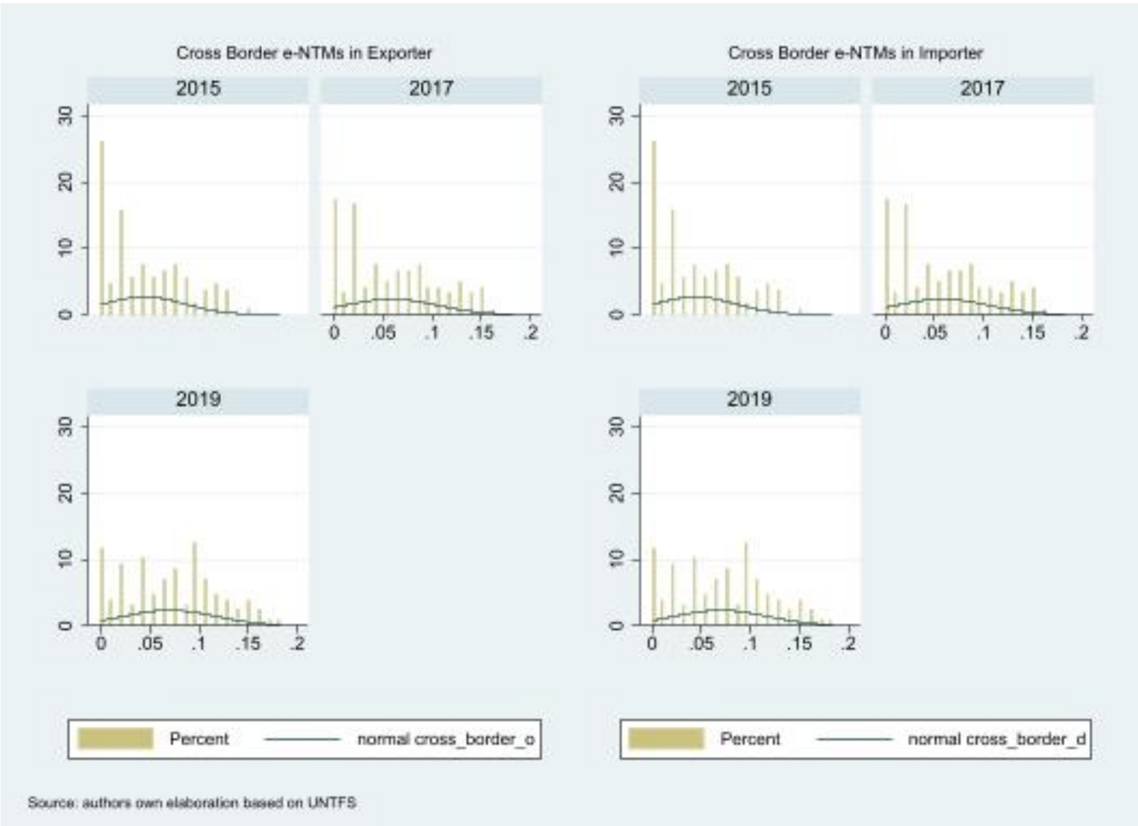


Figure 4.6: Distribution of country-level cross-border NTMs index

In addition to these figures, Table 4.1[56] summarizes the descriptive statistics of the outcome variable and the two main policy variables. It is clear that SSA has much lower agrifood trade flows and lower trade digitalization scores; conversely, Asia has higher agrifood trade flows than the world's average but also lower scores.

Table 4.1: Descriptive statistics of outcome variable and main variables of interest

	(1) Full sample	(2) SSA	(3) SSA mean difference	(4) Asia	(5) Asia mean difference
Bilateral agrifood trade (in current millions USD)	5.74	1.69	-5.02***	6.74	1.23***
	(191.19)	(35.05)	(0.00)	(155.85)	(0.00)
Paperless trade	0.16	0.12	-0.05***	0.14	-0.03***
	(0.08)	(0.05)	(0.00)	(0.09)	(0.00)
Cross-border NTMs paperless trade	0.06	0.03	-0.03***	0.05	-0.02***
	(0.05)	(0.03)	(0.00)	(0.05)	(0.00)
Observations	8,487,830	1,642,752	8,487,830	1,574,304	8,487,830

Means on top and standard deviations in parentheses. Significance levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Mean difference calculated using a t-test.

4.4.2 *Regression results*

This sub-section reports the results following an estimation strategy consisting of a comparison of different specifications of the gravity model from Equation 4.1, taking into account the point estimates and the goodness of fit of each variation. Overall, we observe that: (1) the R-squared varies from 23% to 42% and it is lower for the sub-samples, however, we caveat that simply adding more variables to increase the goodness of fit might have the drawback of compromising internal validity; (2) the gravity variables (GDP and distance) have the signal and the significance expected in the preferred gravity specification; (3) RTAs always have a positive and significant impact on agrifood trade as per trade theory and other studies (Afesorgbor, 2017; Ejones et al., 2021; Fidrmuc, 2004; Mujahid & Kalkuhl, 2016). In sum, the results corroborate

[56] The remaining descriptive statistics of the control and gravity variables are in the Appendix Table 4A.4.

the hypotheses that both standardized electronic documentation and processes (what is referred to as PT) and digitization of NTM documentation (referred to as CB) are important channels to cut transaction costs via facilitation of trade bureaucracy, resulting in higher bilateral trade flows; but this happens in heterogeneous ways. We also find evidence for the role of the policy variables formalities and transparency in facilitating trade, which fits into the conceptual framework – though these other policy variables have more general, and less specific, definitions compared to PT and CB.

Tables 4.2 and 4.3 (and Tables 4A.6-4A.7 in the Appendix) present a summary of the results of the estimations in a three-step order:

a. Sub-divide the policy variables by destination/importer and origin/exporter.

b. Break the tables down into two policy variables, one for PT and the other for CB.

c. Estimate five different model specifications captured by the numbered columns in Tables 4.2 and 4.3: (1-2) PPML with gravity variables and one main policy variable; (3-4) PPML adding socioeconomic controls from Table 4A.3 and one main policy variable; (5-6) PPML with all policy variables and all controls; (7-8, 9-10) the main policy variable (then subsequently all policy variables) with all controls and country-fixed effects employing the PPML with multiple high-dimensional fixed effects (PPML-HDFE) estimator (Correia et al., 2020).

First, I find that PT has a positive and significant impact on bilateral agrifood trade across all specifications, corroborating hypothesis #1. Moreover, I find that both PT and CB have their magnitudes reduced when one moves from the naïve to the more complete models. This suggests that the inclusion of socioeconomic controls and fixed effects absorbs part of the omitted variable bias from the naïve model, which was introduced by unobserved confounders. Variables such as changes in import tariffs or changes in other trade policies not related to participation in trade agreements – which are included in the controls –, and which vary in time, might drive part of this bias. One reason for the inability to control for such confounders is the short time span of the panel that makes it unable to capture new policy shifts. Therefore, I cannot completely rule out some missing correlation between the independent variables and the error term so that the estimates are not fully causally identified. Additionally, I also test models including the two main policy variables (PT and CB) together – see Tables 4.4 (columns 5-9), 4A.8, and 4A.10 (columns 9-10). The results are mostly robust, with the exception of PT at the importer level in the regional sub-samples. However, I caveat that the correlation between the digitalization policy variables is high (see Table 4A.5), which might bias the point estimates, so the preferred specification remains the model with a single policy variable.

The same pattern is observed for all other policy variables, except for institutions, which have a positive effect at the importer level, but a negative effect at the exporter level – despite a very low effect size. This can be interpreted in two possible ways: policies to align trade procedures between customs agencies at the origin have problems either in conception or in implementation; and the policies that governments implement following the passing of legislation matters as well as their effectiveness, which is something this variable does not capture. Lastly, the definition of the measures captured by the institutional variable is rather vague, so I do not put emphasis on its interpretation and omit from the main results.

I highlight three important details about the comparison of magnitudes of the results: (1) the point estimates for PT are greater in magnitude than for CB; (2) the point estimates for PT at the exporter level are larger than for PT at the importer level; (3) the point estimates for CB at the importer level are larger than for CB at the exporter level only in the SSA sub-sample. Two meanings can be attributed for that: first, the impact on agrifood trade of implementing measures related to PT is larger than those related to CB, so LMICs should prioritize general e-facilitation measures; second, changes in e-facilitation tend to have better results when they are implemented by exporters compared to importers, which is a result that bodes well for LMICs, that notoriously lag behind in e-facilitation, as potential exporters, i.e., the measures that yield the best results are under their control.

Next, I turn to CB whose results are in Table 4.3. The direction and the statistical significance of the results are the same compared to those of PT, hence hypothesis #2 is corroborated. This is in line with the conceptual framework as it posits that while imposing NTMs negatively impacts trade, the e-facilitation of NTM documentation stimulates trade as it simplifies the steps and procedures that exporting firms need to follow, thus reducing transaction costs. According to the conceptual framework, the effect at the importer level is more important, since, by definition, NTMs are imposed by the importing country. When comparing the point estimates of CB between the importer and the exporter level, it is not always evident that the effect on importers of CB digitalization is larger than the effect on exporters, apart from the Asian sub-sample.

Regarding the two other policy variables, I find that both the digitalization of trade formalities and transparency have a positive and significant effect on bilateral agrifood trade. Furthermore, the effects are consistent with the previous ones in the sense that the magnitudes are larger for processed products at the exporter level – similar at the importer level. Although the magnitudes of these two other policy variables are, in some cases, larger than those of PT and CB, I put more emphasis in the interpretation of the latter results; the reason is that, following this dissertation's conceptual framework, the definitions of the PT and CB variables carry more relevance and potential for trade facilitation. However, the variables formalities and transparency also have a relevant role in facilitating trade, so they are reported separately, due to the aforementioned high correlation, in the Appendix Tables 4A.6 and 4A.7.

Lastly, I explain in detail how to interpret the coefficients of the regression tables and provide a summary in Figure 4.7. Naturally, since the model takes a log-log form, the results can be directly interpreted in percentage terms from the coefficients, as they capture the partial elasticities:

$$\beta_k = \%\Delta\, Bilateral\ Agrifood\ Trade / \%\Delta\, Digitalization\ Variable\ k. \tag{4.2}$$

Nevertheless, it is not possible to grasp percentage changes without knowing how much a reasonable variation in the digitalization indexes is. For the sake of a plausible interpretation, a 25% increase in the main trade digitalization variables[57] is considered. In the case of PT improvement at the exporter level, agrifood trade increases by 9.7% in the full sample and 11.2% in the SSA; whereas for CB at the importer level, the respective figures are 1.3% and 2.2%; the Asian sub-sample reports even higher coefficients (see Figure 4.7). These results consider the preferred specification, i.e., the PPML model extended with fixed effects and including only one policy variable. These percentage impacts on agrifood trade are non-trivial, considering that implementing trade digitalization is a low-hanging fruit compared to an overhaul in the barriers to trade, or signing new RTAs, or a shift in comparative advantages; therefore, I consider that the results have immediate importance to policymakers. I caveat that increases in the degree of trade digitalization usually happen gradually, so I can only observe small changes within a short period of time, as captured in the descriptive statistics (see Figures 4.4-4.6), therefore these percentage increases are high-bar estimates.

[57] For instance, PT/CB is composed by 10/6 individual measures with implementation of each one scoring between 0-3 points, thus the maximum values equal 0.3/0.18. To compute the percentage change, we consider an increase in one standard deviation as a proportion of the maximum possible value of each index, which is approximately 25% in both cases.

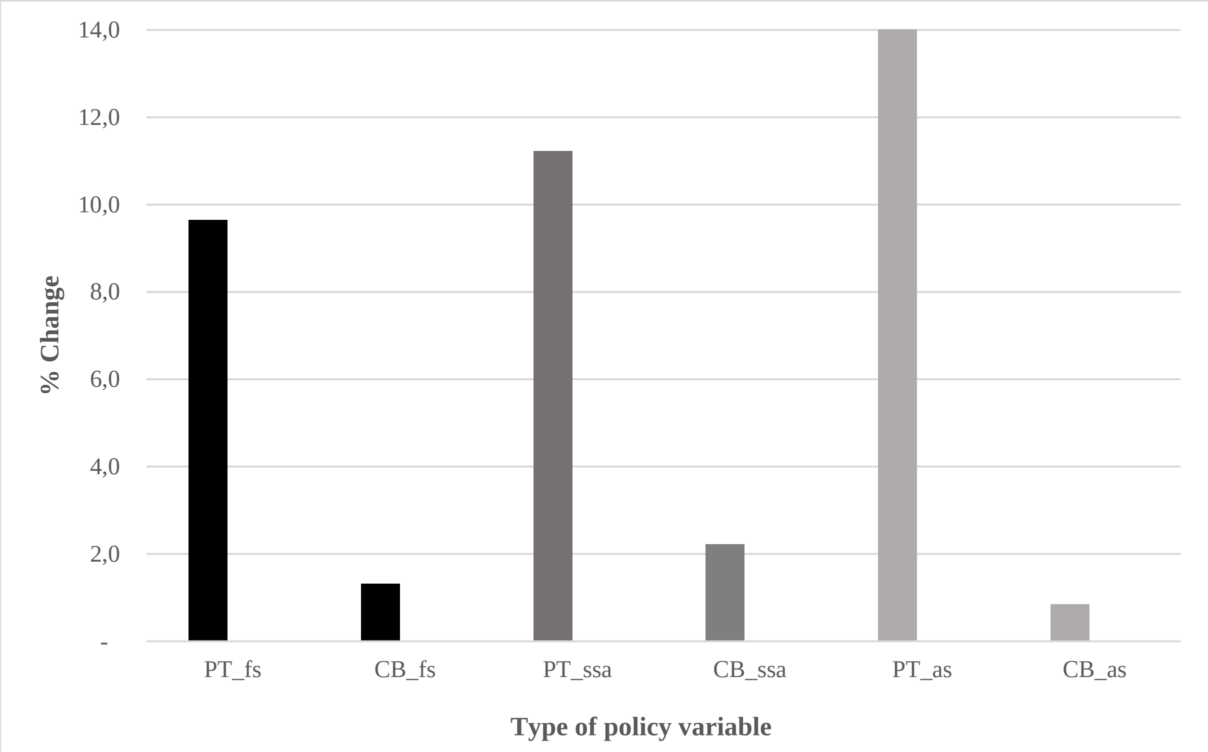


Figure 4.7: Standardized effect of trade digitalization increase on bilateral agri-food trade

No**te:** Policy variables: PT = paperless trade, computed at the exporter level; CB = cross-border non-tariff measures, computed at the importer level. Sub-regions: fs = full sample; ssa = Sub-Saharan Africa; as = Asia. All effects consider a one standard deviation increase in the policy variable adoption. **Source:** author's own elaboration based on secondary data explained in Section 4.3.1.

Table 4.2: Summary of effects of paperless trade digitalization on agrifood trade

Log of Agrifood Trade	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
PT, importers	0.140***		0.108***		0.034***		0.117***		0.040**	
	(0.019)		(0.017)		(0.018)		(0.015)		(0.016)	
PT, exporters		0.449***		0.371***		0.269***		0.386***		0.287***
		(0.021)		(0.020)		(0.022)		(0.019)		(0.021)
Transparency, importers					0.227***				0.217***	
					(0.029)				(0.026)	
Transparency, exporters						0.152***				0.162***
						(0.030)				(0.028)
Formalities, importers					0.119***				0.133***	
					(0.032)				(0.030)	
Formalities, exporters						0.173***				0.156***
						(0.035)				(0.032)
Constant	-9.300***	-8.441***	-9.464***	-8.179***	-8.575***	-7.518***	-2.628***	-5.732**	-1.652**	-5.051***
	(0.829)	(0.872)	(0.195)	(0.194)	(0.212)	(0.215)	(0.673)	(0.532)	(0.670)	(0.537)
Time trend	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Gravity variables	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Socioeconomic Controls	NO	NO	YES	YES	YES	YES	YES	YES	YES	YES
Fixed Effects	NO	NO	NO	NO	NO	NO	YES	YES	YES	YES
Observations	2,978,868	2,978,868	2,739,806	2,739,806	2,739,806	2,739,806	2,618,182	2,618,182	2,618,182	2,618,182
Importing countries	139	180	139	180	139	180	139	180	139	180
Exporting countries	180	139	180	139	180	139	180	139	180	139
Pseudo R^2	0.242	0.237	0.296	0.291	0.298	0.292	0.415	0.384	0.416	0.385
AIC	1.64e+7	1.77e+7	1.49e+7	1.58e+7	1.47e+7	1.58e+7	1.35e+7	1.50e+7	1.35e+7	1.49e+7

Note: Standard errors in parentheses and significance levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Fixed-effects models: R^2 captures the Pseudo-R^2. Policy variables: measured in log.

Table 4.3: Effect of cross-border NTMs digitalization on bilateral agrifood trade: All food groups

Log of Agrifood Trade	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
CB, importers	0.058^{***}		0.045^{***}		0.026^{***}		0.053^{***}		0.029^{***}	
	(0.007)		(0.006)		(0.006)		(0.006)		(0.005)	
CB, exporters		0.107^{***}		0.084^{***}		0.051^{***}		0.088^{***}		0.055^{***}
		(0.008)		(0.007)		(0.006)		(0.006)		(0.006)
Transparency, importers					0.213^{***}				0.203^{***}	
					(0.029)				(0.026)	
Transparency, exporters						0.164^{***}				0.173^{***}
						(0.029)				(0.028)
Formalities, importers					0.132^{***}				0.149^{***}	
					(0.029)				(0.027)	
Formalities, exporters						0.392^{***}				0.387^{***}
						(0.035)				(0.033)
Constant	-9.386^{***}	-9.297^{***}	-9.541^{***}	-9.120^{***}	-8.507^{***}	-7.525^{***}	-2.769^{***}	-6.862^{***}	-1.583^{**}	-5.120^{***}
	(0.817)	(0.851)	(0.185)	(0.182)	(0.213)	(0.217)	(0.671)	(0.529)	(0.670)	(0.538)
Time trend	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Gravity variables	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Socioeconomic Controls	NO	NO	YES	YES	YES	YES	YES	YES	YES	YES
Fixed Effects	NO	NO	NO	NO	NO	NO	YES	YES	YES	YES
Observations	2,978,868	2,978,868	2,739,806	2,739,806	2,739,806	2,739,806	2,618,182	2,618,182	2,618,182	2,618,182
Importing countries	139	180	139	180	139	180	139	180	139	180
Exporting countries	180	139	180	139	180	139	180	139	180	139
Pseudo R^2	0.241	0.234	0.296	0.288	0.298	0.292	0.415	0.381	0.416	0.383
AIC	1.65e+7	1.78e+7	1.50e+7	1.59e+7	1.47e+7	1.58e+7	1.47e+7	1.50e+7	1.47e+7	1.50e+7

Note: Standard errors in parentheses and significance levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Fixed-effects models: R^2 captures the Pseudo-R^2. Policy variables: measured in log.

Table 4.4: Robustness checks of main gravity model

Log of Agrifood Trade	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
PT, importers	0.098***				0.076***		0.088***		
	(0.015)				(0.015)		(0.015)		
PT, exporters		0.347***				0.315***		0.344***	0.366***
		(0.018)				(0.018)		(0.019)	(0.023)
CB, importers			0.044***		0.031***		0.038***		0.017***
			(0.006)		(0.006)		(0.006)		(0.004)
CB, exporters				0.069***		0.029***		0.035***	
				(0.006)		(0.006)		(0.006)	
Sub-Saharan Africa x PT	0.494*** (0.048)	0.611*** (0.022)			0.031*** (0.006)	0.635*** (0.044)			
Sub-Saharan Africa x CB			0.193*** (0.023)	0.192*** (0.010)	0.116*** (0.025)	0.003 (0.020)			
Asia x PT	0.041*	0.306***			0.044*	0.006			
	(0.022)	(0.022)			(0.024)	(0.029)			
Asia x CB			0.132***	0.074***	0.010	0.186***			
			(0.024)	(0.008)	(0.012)	(0.013)			
Constant	-2.236***	-5.409***	-3.102***	-6.830***	-2.012***	-5.319***	-2.432***	-5.626***	-5.667***
	(0.678)	(0.532)	(0.678)	(0.530)	(0.681)	(0.533)	(0.673)	(0.571)	(0.652)
Time trend	YES	YES	YES	YES	YES	YES	YES	YES	YES
Gravity variables	YES	YES	YES	YES	YES	YES	YES	YES	YES
Socioeconomic Controls	YES	YES	YES	YES	YES	YES	YES	YES	YES
Fixed Effects	YES	YES	YES	YES	YES	YES	YES	YES	YES
Observations	2,618,182	2,618,182	2,618,182	2,618,182	2,618,182	2,618,182	2,618,182	2,618,182	1,622,236
Importing countries	139	180	139	180	139	180	139	180	180
Exporting countries	180	139	180	139	180	139	180	139	139
R^2	0.416	0.386	0.415	0.381	0.416	0.387	0.415	0.384	0.377
AIC	1.20e+7	1.20e+7	1.21e+7	1.21e+7	1.39e+7	1.39e+7	1.33e+7	1.33e+7	1.21e+7

Note: Standard errors in parentheses and significance levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Fixed-effects models: R^2 captures the Pseudo-R^2. Policy variables: measured in log. Regional coefficients estimated with interaction terms.

4.4.3 Robustness Checks

Initially, the influence of zero-valued trade is indeed large as the mean of the bilateral trade flow (the dependent variable) is significantly affected when excluding the observations that equal zero, which gives support to the choice of employing the PPML estimator. Then a number of tests are conducted to check for the robustness of the results (Yotov et al., 2016). First, the Akaike information criteria (AIC) compares the trade-off of overfitting vs. underfitting in different models, and it confirms that the preferred specification (PPML-HDFE) has a slightly lower prediction error, thus being the best performing model. Second, I conduct both the Wald test and the likelihood-ratio (LR) test for the joint significance of the regressors in all model specifications, and the tests statistics returned values compatible with a rejection of the null hypothesis that the regressors are jointly equal to zero, both at the destination and at the origin.

The other robustness checks that were conducted are related to the heterogeneity of the effects: (1) regional comparisons for SSA and Asia[58] (see Tables 4A.4, and 4A.8-4A.11); (2) based on the economic sectors classified at the three-digit code by the SITC, the main sample is broken down by the degree of processing of agrifood products into three categories: unprocessed, processed, and highly processed[59].

First, we start with the regional comparisons. When the sample is limited to SSA countries, a significant positive impact of PT on trade is initially documented, but then the magnitude marginally reduces moving to the PPML-HDFE models; a pattern similar to the one observed in the main sample. In the cases of PT and CB at the importer level, I observe a weaker statistical significance in the Asian sub-sample but only when I add additional policy variables, contrasting

[58] Grouped together South Asia and East Asia & Pacific.
[59] The list of all products in each product category is in Appendix Table 4A.1.

with the strong effect of both main policy variables for SSA countries. Focusing on the models with a single policy variable, which are the preferred specifications, I notice that the coefficients of CB at the importer level and PT at the exporter level are 70% and 20% higher in the SSA sub-sample, respectively, whereas in the Asian sub-sample the coefficients are 4% lower and 65% higher, respectively. An alternative way of conducting regional comparisons is by introducing an interaction term between the main policy variables and each region (see Table 4.4, columns 1-6). With this method, the gap between the higher effect of PT at the origin on agrifood trade in Asian countries vis-à-vis SSA countries is inverted; and the higher effect of CB at the destination is widened in SSA countries when compared to Asian ones. It is clear then that the interaction terms for SSA countries have greater magnitudes than Asian.

Contrasting the evidence across all different methods suggests that changes in bureaucratic processes and formalities from a group of LMICs have a larger effect probably because they lag behind in terms of e-bureaucracy implementation, i.e., e-trade facilitation has a higher impact on agrifood trade when it is implemented by these countries as opposed to the destination countries of their exports – although it is not clear if the effect is larger in SSA or in Asia. Moreover, the implementation of digitalization measures related to NTMs compliance by importing countries also has a higher effect in the SSA sub-sample, probably because lowers levels of trade and higher dependence on agrifood trade, as is characteristic of the continent, result in higher marginal increases. It is possible to conclude from this comparison that agrifood trade from the sub-samples of SSA and Asian countries reap higher rewards from trade digitalization than the average, with the exception of CB at the importer level in the Asian sub-sample.

Furthermore, there is the following pattern in the products sub-sample: the effects of PT and CB on bilateral trade of processed agrifood goods is greater than the average, in particular in the regional sub-samples. In terms of PT, the effect at the exporter level is more relevant: the increase in trade of processed and highly processed products is, respectively, 16% and 39% more than the average for each unit change in PT. This pattern is the same in the regional sub-samples. At the importer level, the effect of a CB increase on processed products/unprocessed products is similar/higher to the average impact; this pattern is about the same when analyzing the results using the regional interaction terms. The reason is that the CB policy variable refers to the digitalization of NTMs compliance, and these are concentrated in agricultural products, which are mostly unprocessed. Analogously, the results remain robust when a different specification is estimated, which includes both PT at the exporter level and CB at the importer level – see Table 4.4 (column 9).

I conclude from this set of robustness checks that the main results from the previous sub-section have significant heterogeneities. First, there are greater untapped opportunities for trade facilitation by trade digitalization policies in developing countries. This evidence is important because it corroborates hypothesis #3 that e-trade facilitation has a greater impact on LMICs. Second, SSA and Asian countries also have more opportunities of diversifying its export basket into processed agrifood products and, thus, increasing value addition, since the effect on these goods of general e-trade facilitation is higher than on unprocessed agrifood products.

The last robustness check reported is related to the multilateral resistance term; including all three exporter-year, importer-year, and exporter-importer fixed effects makes the gravity model incur in multicollinearity, thus absorbing all the variation in the policy variables that I am interested in explaining. This happens partly because, in the dataset utilized in this chapter, countries only improve in their trade digitalization indexes – albeit some quite slowly or almost nothing –, so there is either no variation or a small positive variation, but no negative variation. In order to address this potential issue, additional models are estimated with and without the year dummy to check if the results substantially change or not due to the presence of a time trend; the results remain robust[60].

4.5 Conclusions and policy implications

The spread of information and communication technologies unlocks economic potential, in particular by making markets more efficient and reducing price dispersion (Aker, 2010; Aker & Mbiti, 2010). This process is similar in international trade, where technology helps to cut transaction costs, but as with any successful intervention in the agrifood sector, simultaneous complementary interventions on different fronts are needed. The transaction costs considered in this study are related to slow trade facilitation and they often fall on cumbersome procedures at border controls, unfriendly business environments, and misalignments with international trade standards and practices. Facilitating trade through digitalization of such trade and customs procedures can make developing countries attractive commercial destinations. This chapter addresses this issue empirically with a gravity model employed to estimate the effect on bilateral agrifood trade flows of two key policy variables: (1) the introduction of standardized digital documentation and processes, or PT; (2) the introduction of standardized digital documentation and processes related to cross-border non-tariff measures, or CB. In particular, the analysis is enriched by focusing on two sub-regions (Sub-Saharan Africa and Asia) and by documenting the heterogeneity of the impacts by the degree of processing of the agrifood goods that are captured in the trade panel.

[60] These results are available upon request.

The analysis of this chapter measures the impact and trade potential of varying levels of electronic bureaucratic paperwork and barriers to trade by using the data set from the UNTFS, merged and harmonized with two other secondary data sources, i.e., CEPII-BACI and UN Comtrade. The methodology uses the gravity model with fixed effects, which is the workhorse model to explain the impact of changes in certain policy variables on bilateral trade patterns, with the benefit that it is better equipped to deal with long-term and historical determinants of trade as well as it has a high explanatory power. Despite the advantages of the methodology employed, I highlight some limitations: (1) the sample is restricted to three years, so longer panels could identify impacts with different magnitudes; (2) since the estimation methodology adopted here is not a general equilibrium model, I emphasize that the results account only for partial equilibrium effects; (3) although the model controls for fixed effects there might still be unobserved time-variant confounders, so Iremain cautious about identification and prefer to talk about associations; (4) the research question analyzed in this chapter deals only with the benefits of implementing trade digitalization and not with the costs, hence there is a need for future studies on cost-benefit analysis.

The findings point to large and significant benefits of trade bureaucracy digitalization on bilateral agrifood trade, with important heterogeneities. First, while the main measures to digitize trade bureaucracy that are analyzed (PT and CB) increase agrifood trade in the cross-country sample, the trade institutions variable of interest does not have the same clear effect. Second, this effect I identify for the whole sample is even stronger in the case of SSA and Asia. The results are robust to the inclusion of time- and unobserved country-fixed effects. Third, once the sample is broken down by degree of processing of the agrifood products, I observe that the positive effects are also stronger for processed products at the exporter level, which bodes well for an industrial policy focusing on adding value to local raw materials. This could happen because processed products tend to be more complex than unprocessed products in terms of their composition – with a higher degree of imported inputs and parts and a more complex value chain –, thus requiring more detailed and a larger volume of documentation to clear customs procedures and rules of origin. Last, I observe that digitalization measures at the importer level (destination) that facilitate compliance with NTMs also play an important role in stimulating agrifood exports.

The following is highlighted from the results: (1) PT measures have a higher impact when they are implemented in exporting countries, so this is an important finding for LMICs to facilitate electronic trade bureaucracy; (2) this pattern is stronger in the Asian and SSA cases, demanding African policymakers to continue and deepen their efforts of facilitating trade via digitalization of bureaucratic procedures; (3) simply passing legislation to empower customs authorities and to create legal frameworks is not enough to foster trade, but the actual content of legislation as well as the effectiveness and implementation of trade institutions are crucial; (4) the costs and technical requirements for implementing trade facilitation via digitalization of customs procedures, especially for LMICs, need to be considered.

Finally, I highlight two policy implications emerging from the analysis and one opportunity for further research. In general, the findings support the use of aid-for-trade programs for trade facilitation (Pettersson & Johansson, 2013), particularly if focused on improving digitalization measures. And although the COVID-19 pandemic led countries to retrench into domestic markets, the global and regional trade of agrifood products was less directly blocked as countries realized it is a lifeline to ensure domestic food security (Liverpool-Tasie et al., 2021). Therefore, and especially considering that RTAs increase agrifood trade (Mujahid & Kalkuhl, 2016), the momentum to deepen regional trade integration with the inception of the AfCFTA should be seized by SSA countries, which should take this opportunity to improve the digitalization of trade procedures and bureaucracies in order to increase value addition and advance their position in agrifood trade markets. An interesting line of future research, that requires granular input-output data, is to breakdown the effects of trade and customs digitalization by production steps and degree of participation in GVCs (considering both backward and forward linkages), since the gains associated with trade tend to be higher when firms and countries are connected to GVCs in comparison to when they engage in simple final products trade (Antràs & de Gortari, 2020).

CHAPTER 5

General conclusions and policy implications

The concluding chapter of this dissertation has three objectives: (1) summarize the main results of the analytical chapters focusing on the interconnections between the findings; (2) highlight the main takeaways from the concluding sections of each chapter and propose a set of policy recommendations, based on actionable empirical evidence; (3) acknowledge the limitations of this study and use them as a starting point for suggested lines of future research.

5.1 Summary of the main results

As explained in the sub-section that defines key concepts (Section 1.3), reductions in transaction costs in agricultural value chains are essential to increase agricultural productivity, in particular in the small farm sector, which, in turn, is a pre-condition for labor to shift to higher productivity economic activities, thus fueling structural transformation (Johnston & Mellor, 1961). The importance of the work undertaken in this dissertation is that it provides empirical evidence from different contexts on how better market access and improved market connections are essential to reduce transaction costs faced by SSFs in the Tanzanian context. This is set against the backdrop of the ongoing transformation of the agrifood system that can catalyze the achievement of the goal of inclusive growth and prosperity.

Chapter 2, the first analytical chapter, analyzed the welfare effects of different agricultural marketing channels in which maize SSFs operate in Tanzania. The variable of interest – vertical market linkages – captures these channels and entails market connections that cut transaction costs, which could be a solution to the pervasive problem of high transaction costs that SSFs face. The chapter employed panel data methods and found that SSFs that have direct connections with AP firms and market arrangements similar to CF present better welfare outcomes when compared to SSFs with connections to cooperatives and local merchants. Effects on welfare, as measured by household income and food security, are also particularly salient.

Chapter 3 extends the analysis of the previous chapter by focusing on the welfare effects of an emerging and more sophisticated type of marketing channel, i.e., CF, in the Tanzanian cashew nuts sector, the most important cash crop of the country. The CF arrangements were agreed between SSFs and cashew processors in simple terms, setting only prices, quantities, and a rough quality benchmark. The logic is similar to the previous chapter: CF reduces transaction costs related to market and price risks through direct connections between farmers and processing firms. When compared to farmers that used exclusively the government-run auction system, which is the control group, SSFs who engaged with CF increased their cashew output and reduced food insecurity.

The fourth and last analytical chapter turned the focus from farming household surveys to international trade whereby it investigated the effects of digitalization of trade and customs processes and bureaucracies on bilateral agrifood trade. These digitalization measures cut transaction costs and, thus, facilitate agrifood trade. The effects are particularly salient for processed agrifood products and for two regions: SSA and Asia.

The first two analytical chapters found reductions on food insecurity associated with participation in agricultural market linkages – the effect from CF participants in the third chapter was larger than in the second chapter. The main difference between the variables of interest of these two chapters and, thus, of the effects analyzed, is that Chapter 2 considers a more comprehensive set of agricultural market linkages whilst Chapter 3 focuses solely on CF. Thus, the positive results of Chapter 2 are corroborated by a more specific case study, but with a focus on a cash crop (cashew nuts) as opposed to a marketable food crop (maize). The transmission mechanisms of these effects are lower market risk and price risk. SSFs connecting to VMLs reduce their risks of commercialization and increase their output prices. Furthermore, besides receiving higher output prices, CF participants reduce risk compared to the auction participants because the auction system has frequent and inherent price fluctuations; apart from recurrent rejection of bids deemed too low by cooperative leaders. The comparison between the risk reduction that SSFs attain in the two chapters applies especially to VML2 because this variable of interest is built considering CF schemes and other marketing linkages that involve direct connections between SSFs and AP firms. And these risk reductions are even more pronounced compared to subsistence farmers who have only own-consumption[61] as marketing channel. In sum, the combination of these two studies advances scholarly knowledge on in which countries, for which crops, and with which connections to agricultural market linkages SSFs' can improve household welfare.

And, lastly, the breadth of the analysis covered in this dissertation is enhanced with the study of a different type of transaction cost reduction in Chapter 4: digitalization measures that increase agrifood trade. Since the export of agrifood products is generally done by AP firms, it is possible to conclude that higher agrifood trade is a potential catalyzer of opportunities for more frequent and higher quality connections between SSFs and AP firms. For instance, CF schemes and connection to the processors' VML could contribute to Tanzania's agrifood exporting capacity, so increased agrifood trade could be a transmission channel to higher rural income. Indirectly, the results of this last analytical chapter corroborate the importance of the findings of the two previous

[61] Even though, in theory, subsistence farmers would not incur price risk precisely because they consume what they produce, they still incur price risk because implicitly they forego revenues from selling their output.

chapters in the sense that improving market access and reducing barriers to trade can increase rural household welfare and promote the AP industry.

5.2 Policy implications

Based on the empirical findings of the analytical chapters, I list below priority areas of intervention for effective public policies with potential to increase agricultural productivity, support rural income, and improve food security in the context of Tanzania.

First, improvement of agricultural productivity is paramount because of its potential impact on poverty reduction and food insecurity amelioration in rural areas, as demonstrated in Chapters 2 and 3, and on wider industrialization efforts via the AP channel, as shown in Chapter 4. Here, increases in productivity encompass both technological improvements and changes in the products mix towards higher value crops or other products with higher value added. There are three important policy instruments to increase agricultural productivity: (1) expansion of agricultural extension services (Anderson & Feder, 2004; Buehren et al., 2019); (2) higher quality targeted advice from extension services for specific crops and regions (Antwi-Agyei & Stringer, 2021; Ragasa & Mazunda, 2018); (3) investments in farmers' training and improvements in the quality of the vocational educational system (Kabasa et al., 2015). To these ends, investments via ministerial budget allocation and/or parastatal activities should be directed towards better targeted agricultural extension services, professional and vocational training in the AP industry, seed capital from the Tanzania Agricultural Development Bank (TADB) for promising business plans, and improved access to basic utility services. Notably, the GoT's spending in agriculture and rural development (about 2.6%[62]) is well under both its pledge[63] to spend 10% of the budget in the sector and the SSA average disbursement (Pernechele et al., 2021).

Second, agricultural production suffers from limited availability and high costs of inputs, poor communication and transportation links, and other constraints to output commercialization discussed in Chapter 2. The lack of reliable access to markets reduces demand for investments that could increase both agricultural production and productivity, thus hampering the ongoing process of structural transformation. The policy instruments to address these issues are the promotion of CF schemes and better coordination of crop marketing boards with farmers and processors because they stimulate the engagement of the private sector with SSFs thereby increasing the efficiency of rural markets. With this in mind, policymakers from the agricultural sector – Ministry of Agriculture (MoA), crop boards, and parastatals – should focus on improving coordination between all stakeholders and increase information sharing to reduce information asymmetries. Lack of coordination often leads to policy incoherence as well as public policies formulated with unclear targets and infeasible objectives. It is important that government officials set feasible objectives rather than unrealistic ones, so that buy-in across government agencies and society can be secured. One example from the fieldwork conducted for Chapter 3 is regarding the RCN output targets set by the MoA. Since they are often infeasible and not jointly discussed, the CBT cannot appropriately implement its policies on-the-ground alongside farmers and AP firms. Another example is related to the mandatory auction system that the CBT uses for RCN sales: as documented in the third chapter, this system delivers lower welfare to SSFs vis-à-vis a simple CF marketing scheme.

Third, the manifold risks that SSFs face should be reduced. The main types of risk discussed in this

[62] Budget Speech 2022/23: https://www.mof.go.tz/uploads/documents/en-1655216926-BUDGET%20SPEECH%202022-23%20ENGLISH%20VERSION.pdf <Accessed on 10 November 2023>.

[63] Comprehensive Africa Agriculture Development Programme (CAADP) from the African Union (AU).

dissertation are related to crop prices and crop marketing. The first type of risk is simply the risk that SSFs face of receiving prices that are too low and/or large price fluctuations. The other risk is associated with difficulties in finding buyers for their crops and the costs associated with this search as well as monitoring and enforcement of formal contracts or informal relationships. Risk reduction also helps to reduce vulnerability and strengthen resilience to covariate shocks, which is essential to deal with increasingly severe climatic risks as well as economic risks. These are topics discussed in Chapters 2 and 3: reducing such risks stimulates rural incomes and, consequently, improves household welfare. The policy instruments to achieve that are subsidies to the financing of crops and inputs from the TADB, promotion of a gradual shift from cash crops auction systems to CF, and assistance to smallholders in finding markets.

Lastly, not all policies have positive effects for all stakeholders because some of them are zero-sum, i.e., they create winners and losers. Tanzanian policymakers tend to prefer inefficient policies with short-term gains for their constituency – mostly farming households – over efficient policies bringing long-term gains for the majority of stakeholders in the value chain (Msami & Wangwe, 2016). This demands policymakers to carefully consider the implications of their actions and the interests involved, aiming to maximize aggregate social welfare. For instance, the policy preferences of rural producers likely differ from the preferences of urban consumers as well as the interests of processors and farmers are not aligned at all times. Since winners, who accrue most of the benefits, are unlikely to (fully) compensate losers, interventions to reform or build value chains are not neutral activities. Some policies will be supported by certain social groups, likely those who benefit most, and resisted by others, likely those who benefit less or even lose. Chapter 3 has an example of this: due to political economy considerations, the government often intervenes in cashew auctions to support farmers with price floors that processors consider too high.

5.3 Limitations and suggestions for future research

This last sub-section explores the limitations of this dissertation to set an agenda for future research. The limitations are related to external validity, absence of an explicit consideration of prices in the conceptual framework, issues that cannot be analyzed using the data sets employed, and the role of environmental and sustainability considerations.

First, although some of the socioeconomic conditions analyzed in the case of Tanzania might be analogous to similar contexts, I acknowledge limitations in the external validity of the findings from Chapters 2 and 3, since both utilized data sets from a single country. An interesting agenda for future research would be on different definitions of agricultural market linkages and on how farming households use mixed channels as profit and/or utility-maximizing strategies. Relatedly, there are opportunities for future research on the effects that these linkages have not only on SSFs welfare but also as catalyzers of the AP industry due to the ability of such linkages in connecting farmers with processors. An interesting starting point for this research agenda would be to conduct further studies using data from other countries contained in the LSMS-ISA project.

In a similar way, I remain cautious about extrapolating the findings from Chapter 3, in which the primary data utilized comes from the author’s own fieldwork. The reason is that the chapter focuses on one particular cash crop (cashew nuts) in a particular country (Tanzania). Hence, one immediate opportunity for future research would be to either expand the time span of this data set (through the creation of a panel), or to expand the scope of the analysis, i.e., shifting the focus to a different country and/or cash crop-CF scheme combination. This could be a fruitful endeavor as there is currently a lack of rigorous impact analyses about the effects of CF in the cashew nuts sector in general and in Tanzania, in specific – and also studies employing panel data are rare. Another interesting path for future research on the CF topic would be to investigate the determinants of adoption of different types of contracts as well as what are the benefits for SSFs of adopting these different types of contracts (Dubbert & Abdulai, 2021; Ruml & Qaim, 2020, 2021b). This research could be done on the same country-crop setting or in different contexts.

Notwithstanding the caveats about limitations on external validity, I emphasize that the key socioeconomic and agronomic characteristics of the Tanzanian smallholder farming sector are similar to those found in other LDCs, especially in SSA. This means that some of the results, findings, and policy recommendations of the first two analytical chapters are relevant beyond the specific cases studied, hence these findings can inform future research along similar lines. Consequently, it is important that researchers and practitioners adopt a perspective that considers the local reality and the specific demands of local communities when using the findings and policy recommendations of this thesis to conduct research in other contexts.

Regarding Chapter 4, the most immediate avenue for future research is to expand the time span of the panel as new data is published and/or focus the analysis in regions other than SSA or Asia. Another fruitful opportunity for further research would be to expand the span of products considered and compare the effects between industrial and agrifood goods. While the gravity model is the workhorse methodology to analyze the effect of changes in specific policies on bilateral trade, it captures neither backward nor forward linkages. These would allow the researcher to draw conclusions on which specific industrial sector has more impact on jobs and growth, or which sector benefits more from digitalization of trade and customs procedures. An interesting path for future research is, therefore, to employ multi-region input-output models.

One limitation of this dissertation is an explicit consideration of the causes and consequences of price fluctuations and price changes, which would enable the precise determination of efficiency and equity effects. This happened because the data sets employed either do not cover prices at each point of the value chain (e.g., farmgate, local traders, AP firms, local markets, etc.) or were not able to retrieve granular price data due to technical limitations. Prices are the main mechanism that clears supply and demand in goods markets and, hence, provide key information for producers and consumers to adjust their actions. Despite this limitation, the conceptual framework captures price movements indirectly regarding changes in transaction costs since they are an important component of goods' prices, especially tradeable agrifood goods. Thus, the potential consequences from changes in transaction costs differ depending on which stage of the value chain or which market actor is considered. For instance, the introduction of contractual arrangements between SSFs and processors likely reduces market risk, which translates into lower transaction costs and, consequently, higher output prices.

The data sets employed have two other methodological limitations. First, they are constrained to the household-level and, thus, lack individual-level information, since going onto an individual level was beyond the scope of this thesis. Due to local and regional socioeconomic dynamics, such as gender and ethnic power structures, the distribution of income, food, and decision-making inside households is not even across all members, rather often concentrated in the male household head. Hence, it is not possible to either analyze or draw conclusions about individual-specific effects and intra-household distribution of resources, which is a fruitful avenue for future research. And second, the data sets contain observational data, thus prone to self-selection bias. Appropriate empirical methodologies were employed to address this issue – such as fixed effects and CRE panel models, matching techniques, and an instrumental variable approach, but they still have shortcomings with respect to selection on unobserved time-varying covariates. A definitive solution to this issue would be to have conducted randomized controlled trials (RCTs). Investing in RCTs would be an impactful future research agenda, although it was beyond the scope of this thesis. The reason for this is that RCTs are rare in the literature of CF and agricultural market linkages for clear reasons: because of high costs and complexity of implementation and because of the intricate nature of the research questions that these literatures analyze (e.g., it is not an easy task to disentangle causes and effects).

Last but not least, although social, economic, and agronomic factors were considered, environmental and sustainability concerns were beyond the scope of this thesis and, thus, were not directly addressed. In this regard, the following research questions could guide an interesting agenda of future research in this area:

- How can sustainable intensification drive the modernization of agricultural market linkages in the Global South?
- How can the connection of SSFs into agricultural market linkages (especially CF) increase rural households' welfare while integrating a nature-positive production framework?

Overall, this thesis has contributed to the literature of agricultural marketing by providing original empirical evidence showing that cutting transaction costs is key to improve the connection of Tanzanian SSFs into agricultural market linkages. It also studied how transaction costs in the form of digitalization of customs measures can increase agrifood trade and, thereby, stimulate connections between exporting processors and farmers. Finally, based on these empirical evidence from the analytical chapters, this dissertation has also provided actionable and evidence-based recommendations to better inform policymakers.

6 References

Abadie, A., Athey, S., Imbens, G. W., & Wooldridge, J. M. (2022). When Should You Adjust Standard Errors for Clustering?*. *The Quarterly Journal of Economics*, qjac038. https://doi.org/10.1093/qje/qjac038

Abdulai, A., & Aubert, D. (2004). A cross-section analysis of household demand for food and nutrients in Tanzania. *Agricultural Economics*, *31*(1), 67–79. https://doi.org/10.1016/j.agecon.2003.03.001

Achterbosch, T., van Berkum, S., & Meijerink, G. W. (2014). *Cash crops and food security; contributions to income, livelihood risk and agricultural innovation* (LEI Report). Wageningen University.

Adabe, K. E., Abbey, A. G., Egyir, I. S., Kuwornu, J. K. M., & Anim-Somuah, H. (2019). Impact of contract farming on product quality upgrading: The case of rice in Togo. *Journal of Agribusiness in Developing and Emerging Economies*, *9*(4), 314–332. https://doi.org/10.1108/JADEE-04-2018-0051

Adamopoulos, T., & Restuccia, D. (2022). Geography and Agricultural Productivity: Cross-Country Evidence from Micro Plot-Level Data. *The Review of Economic Studies*, *89*(4), 1629–1653. https://doi.org/10.1093/restud/rdab059

AfDB. (2019). *African Economic Outlook 2019. Macroeconomic Performance and Prospects. Jobs, Growth and Firm Dynamism. Integration for Africa's Prosperity.* (African Economic Outlook). African Development Bank. https://www.afdb.org/fileadmin/uploads/afdb/Documents/Publications/2019AEO/AEO_2019-EN.pdf

Afesorgbor, S. K. (2017). Revisiting the effect of regional integration on African trade: Evidence from meta-analysis and gravity model. *The Journal of International Trade & Economic Development*, *26*(2), 133–153. https://doi.org/10.1080/09638199.2016.1219381

Ajwang, F. (2020). Relational Contracts and Smallholder Farmers' Entry, Stay and Exit, in Kenyan Fresh Fruits and Vegetables Export Value Chain. *The Journal of Development Studies*, *56*(4), 782–797. https://doi.org/10.1080/00220388.2019.1618451

Aker, J. C. (2010). Information from Markets Near and Far: Mobile Phones and Agricultural Markets in Niger. *American Economic Journal: Applied Economics*, *2*(3), 46–59. https://doi.org/10.1257/app.2.3.46

Aker, J. C., & Mbiti, I. M. (2010). Mobile Phones and Economic Development in Africa. *Journal of Economic Perspectives*, *24*(3), 207–232. https://doi.org/10.1257/jep.24.3.207

Ameye, H., De Weerdt, J., & Gibson, J. (2021). Measuring macro- and micronutrient consumption in multi-purpose surveys: Evidence from a survey experiment in Tanzania. *Food Policy*, *102*, 102042. https://doi.org/10.1016/j.foodpol.2021.102042

Anderson, J. E. (1979). A Theoretical Foundation for the Gravity Equation. *The American Economic Review*, *69*(1), 106–116.

Anderson, J. E., & van Wincoop, E. (2003). Gravity with Gravitas: A Solution to the Border Puzzle. *American Economic Review*, *93*(1), 170–192. https://doi.org/10.1257/000282803321455214

Anderson, J. E., & van Wincoop, E. (2004). Trade Costs. *Journal of Economic Literature*, *42*(3), 691–751. https://doi.org/10.1257/0022051042177649

Anderson, J. R., & Feder, G. (2004). Agricultural Extension: Good Intentions and Hard Realities. *The World Bank Research Observer*, *19*(1), 41–60. https://doi.org/10.1093/wbro/lkh013

Angrist, J., & Pischke, J.-S. (2009). *Mostly Harmless Econometrics* (1st Edition). Princeton University Press. https://press.princeton.edu/books/paperback/9780691120355/mostly-harmless-econometrics

ANSAF. (2022). *Opportunity costs of not processing cashew domestically and the loss from exporting raw*. Agricultural Non-State Actors Forum (ANSAF).

Antràs, P., & de Gortari, A. (2020). On the Geography of Global Value Chains. *Econometrica*, *88*(4), 1553–1598. https://doi.org/10.3982/ECTA15362

Antwi-Agyei, P., & Stringer, L. C. (2021). Improving the effectiveness of agricultural extension services in supporting farmers to adapt to climate change: Insights from northeastern Ghana. *Climate Risk Management*, *32*(100304). https://doi.org/10.1016/j.crm.2021.100304

Arouna, A., Michler, J. D., & Lokossou, J. C. (2021). Contract farming and rural transformation: Evidence from a field experiment in Benin. *Journal of Development Economics*, *151*, 102626. https://doi.org/10.1016/j.jdeveco.2021.102626

Arvis, J.-F., Duval, Y., Shepherd, B., Utoktham, C., & Raj, A. (2016). Trade Costs in the Developing World: 1996–2010. *World Trade Review*, *15*(3), 451–474. https://doi.org/10.1017/S147474561500052X

Arvis, J.-F., & Shepherd, B. (2013). The Poisson quasi-maximum likelihood estimator: A solution to the 'adding up' problem in gravity models. *Applied Economics Letters*, *20*(6), 515–519. https://doi.org/10.1080/13504851.2012.718052

Asfaw, S., Kassie, M., Simtowe, F., & Lipper, L. (2012). Poverty Reduction Effects of Agricultural Technology Adoption: A Micro-evidence from Rural Tanzania. *The Journal of Development Studies*, *48*(9), 1288–1305. https://doi.org/10.1080/00220388.2012.671475

Avery, E., Baragwanath, T., Deuss, A., Marlet, E., Sorescu, S., & van Tongeren, F. (2021). *Digital Opportunities for Sanitary and Phytosanitary (SPS) Systems and the Trade Facilitation Effects of SPS Electronic Certification* (Working Paper No. 152; OECD Food, Agriculture and Fisheries Paper). OECD.

Barrett, C. B. (2008). Smallholder market participation: Concepts and evidence from eastern and southern Africa. *Food Policy*, *33*(4), 299–317. https://doi.org/10.1016/j.foodpol.2007.10.005

Barrett, C. B. (2010). Measuring Food Insecurity. *Science*, *327*(5967), 825–828. https://doi.org/10.1126/science.1182768

Barrett, C. B., Bachke, M. E., Bellemare, M. F., Michelson, H. C., Narayanan, S., & Walker, T. F. (2012). Smallholder Participation in Contract Farming: Comparative Evidence from Five Countries. *World Development*, *40*(4), 715–730. https://doi.org/10.1016/j.worlddev.2011.09.006

Barrett, C. B., Christiaensen, L., Sheahan, M., & Shimeles, A. (2017). On the Structural Transformation of Rural Africa. *Journal of African Economies*, *26*(suppl_1), i11–i35. https://doi.org/10.1093/jae/ejx009

Barrett, C. B., Christian, P., & Shimeles, A. (2018). The processes of structural transformation of African agriculture and rural spaces. *World Development*, *105*, 283–285. https://doi.org/10.1016/j.worlddev.2018.02.019

Barrett, C. B., Reardon, T., Swinnen, J., & Zilberman, D. (2022). Agri-food Value Chain Revolutions in Low- and Middle-Income Countries. *Journal of Economic Literature*, *60*(4), 1316–1377. https://doi.org/10.1257/jel.20201539

Bellemare, M. F. (2012). As You Sow, So Shall You Reap: The Welfare Impacts of Contract Farming. *World Development*, *40*(7), 1418–1434. https://doi.org/10.1016/j.worlddev.2011.12.008

Bellemare, M. F., Barrett, C. B., & Just, D. R. (2013). The Welfare Impacts of Commodity Price Volatility: Evidence from Rural Ethiopia. *American Journal of Agricultural Economics*, *95*(4), 877–899. https://doi.org/10.1093/ajae/aat018

Bellemare, M. F., & Bloem, J. R. (2018). Does contract farming improve welfare? A review. *World Development*, *112*, 259–271. https://doi.org/10.1016/j.worlddev.2018.08.018

Bellemare, M. F., Bloem, J. R., & Lim, S. (2022). Producers, consumers, and value chains in low- and middle-income countries. In *Handbook of Agricultural Economics*. Elsevier. https://doi.org/10.1016/bs.hesagr.2022.03.005

Bellemare, M. F., Lee, Y. N., & Novak, L. (2021). Contract farming as partial insurance. *World Development*, *140*, 105274. https://doi.org/10.1016/j.worlddev.2020.105274

Bellemare, M. F., & Lim, S. (2018). In All Shapes and Colors: Varieties of Contract Farming. *Applied Economic Perspectives and Policy*, *40*(3), 379–401. https://doi.org/10.1093/aepp/ppy019

Bellemare, M. F., & Novak, L. (2017). Contract Farming and Food Security. *American Journal of Agricultural Economics*, *99*(2), 357–378. https://doi.org/10.1093/ajae/aaw053

Benali, M., Brümmer, B., & Afari-Sefa, V. (2018). Smallholder participation in vegetable exports and age-disaggregated labor allocation in Northern Tanzania. *Agricultural Economics*, *49*(5), 549–562. https://doi.org/10.1111/agec.12441

Bennett, M. K. (1941). Wheat in national diets. *Wheat Studies*, *18*, 37–76.

Berg, E. (2013). Are poor people credit-constrained or myopic? Evidence from a South African panel. *Journal of Development Economics*, *101*, 195–205. https://doi.org/10.1016/j.jdeveco.2012.10.002

Bergstrand, J. H. (1985). The Gravity Equation in International Trade: Some Microeconomic Foundations and Empirical Evidence. *The Review of Economics and Statistics*, *67*(3), 474–481. https://doi.org/10.2307/1925976

Biggeri, M., Burchi, F., Ciani, F., & Herrmann, R. (2018). Linking small-scale farmers to the durum wheat value chain in Ethiopia: Assessing the effects on production and wellbeing. *Food Policy*, *79*, 77–91. https://doi.org/10.1016/j.foodpol.2018.06.001

Binswanger, H. P., & Rosenzweig, M. R. (1986). Behavioural and material determinants of production relations in agriculture. *The Journal of Development Studies*, *22*(3), 503–539. https://doi.org/10.1080/00220388608421994

Blandon, J., Henson, S., & Islam, T. (2010). The Importance of Assessing Marketing Preferences of Small-scale Farmers: A Latent Segment Approach. *The European Journal of Development Research*, *22*(4), 494–509. https://doi.org/10.1057/ejdr.2010.26

Bouët, A., Tadesse, G., & Zaki, C. (2021). *Africa Agriculture Trade Monitor 2021*. AKADEMIYA2063 & IFPRI. https://doi.org/10.54067/9781737916406

Bratt, M. (2017). Estimating the bilateral impact of nontariff measures on trade. *Review of International Economics*, *25*(5), 1105–1129. https://doi.org/10.1111/roie.12297

Brück, T., d'Errico, M., & Pietrelli, R. (2019). The effects of violent conflict on household resilience and food security: Evidence from the 2014 Gaza conflict. *World Development*, *119*, 203–223. https://doi.org/10.1016/j.worlddev.2018.05.008

Brun, J.-F., Carrère, C., Guillaumont, P., & de Melo, J. (2005). Has Distance Died? Evidence from a Panel Gravity Model. *The World Bank Economic Review*, *19*(1), 99–120. https://doi.org/10.1093/wber/lhi004

Buehren, N., Goldstein, M., Molina, E., & Vaillant, J. (2019). The impact of strengthening agricultural extension services on women farmers: Evidence from Ethiopia. *Agricultural Economics*, *50*(4), 407–419. https://doi.org/10.1111/agec.12499

Burger, M., Oort, F. van, & Linders, G.-J. (2009). On the Specification of the Gravity Model of Trade: Zeros, Excess Zeros and Zero-inflated Estimation. *Spatial Economic Analysis*, *4*(2), 167–190. https://doi.org/10.1080/17421770902834327

Cahyadi, E. R., & Waibel, H. (2016). Contract Farming and Vulnerability to Poverty among Oil Palm Smallholders in Indonesia. *The Journal of Development Studies*, *52*(5), 681–695. https://doi.org/10.1080/00220388.2015.1098627

Caliendo, M., & Kopeinig, S. (2008). Some Practical Guidance for the Implementation of Propensity Score Matching. *Journal of Economic Surveys*, *22*(1), 31–72. https://doi.org/10.1111/j.1467-6419.2007.00527.x

Campenhout, B. V., Minten, B., & Swinnen, J. F. M. (2021). Leading the way – foreign direct investment and dairy value chain upgrading in Uganda. *Agricultural Economics*, *52*(4), 607–631. https://doi.org/10.1111/agec.12638

Carletto, C., Corral, P., & Guelfi, A. (2017). Agricultural commercialization and nutrition revisited: Empirical evidence from three African countries. *Food Policy*, *67*, 106–118. https://doi.org/10.1016/j.foodpol.2016.09.020

Casaburi, L., & Willis, J. (2018). Time versus State in Insurance: Experimental Evidence from Contract Farming in Kenya. *American Economic Review*, *108*(12), 3778–3813. https://doi.org/10.1257/aer.20171526

Cerulli, G. (2015). Methods Based on Selection on Unobservables. In G. Cerulli (Ed.), *Econometric Evaluation of Socio-Economic Programs: Theory and Applications* (pp. 161–227). Springer Berlin. https://doi.org/10.1007/978-3-662-46405-2_3

CFS. (2012). *Global Strategic Framework for Food Security and Nutrition* (Conference Report CFS 2012/39/5; p. 39). Committee on World Food Security, Food and Agriculture Organization (FAO). http://www.fao.org/fileadmin/user_upload/bodies/CFS_sessions/39th_Session/39emerg/ME498E_CFS_2012_39_5_Add_1_Rev_1.pdf

Chege, C. G. K., Andersson, C. I. M., & Qaim, M. (2015). Impacts of Supermarkets on Farm Household Nutrition in Kenya. *World Development*, *72*, 394–407. https://doi.org/10.1016/j.worlddev.2015.03.016

Cheng, I.-H., & Wall, H. J. (2005). Controlling for Heterogeneity in Gravity Models of Trade and Integration. *Federal Reserve Bank of St. Louis Review*, *87*(1), 49–63. https://doi.org/10.20955/r.87.49-64

Christiaensen, L., De Weerdt, J., & Todo, Y. (2013). Urbanization and poverty reduction: The role of rural diversification and secondary towns1. *Agricultural Economics*, *44*(4–5), 435–447. https://doi.org/10.1111/agec.12028

Christiaensen, L., Demery, L., & Kuhl, J. (2011). The (evolving) role of agriculture in poverty reduction—An empirical perspective. *Journal of Development Economics*, *96*(2), 239–254. https://doi.org/10.1016/j.jdeveco.2010.10.006

Cipollina, M., De Benedictis, L., Salvatici, L., & Vicarelli, C. (2016). *Policy Measurement and Multilateral Resistance in Gravity Models* (Working Papers LuissLab No. 16130). Dipartimento di Economia e Finanza, LUISS Guido Carli. https://econpapers.repec.org/paper/luilleewp/16130.htm

Civelek, M. E., Çemberci, M., Uca, N., Çelebi, Ü., & Özalp, A. (2017). Challenges of Paperless Trade: Redesign of the Foreign Trade Processes and Bundling Functions of Traditional

Documents. *International Business Research*, *10*(2), 74. https://doi.org/10.5539/ibr.v10n2p74

Coase, R. H. (1937). The Nature of the Firm. *Economica*, *4*(16), 386–405. JSTOR. https://doi.org/10.2307/2626876

Coates, J., Swindale, A., & Bilinsky, P. (2007). *Household Food Insecurity Access Scale (HFIAS) for Measurement of Food Access: Indicator Guide:* (Food and Nutrition Technical Assistance (FANTA III project)). USAID. https://doi.org/10.1037/e576842013-001

Collier, P., & Dercon, S. (2014). African Agriculture in 50 Years: Smallholders in a Rapidly Changing World? *World Development*, *63*, 92–101. https://doi.org/10.1016/j.worlddev.2013.10.001

Correia, S., Guimarães, P., & Zylkin, T. (2020). Fast Poisson estimation with high-dimensional fixed effects. *The Stata Journal*, *20*(1), 95–115. https://doi.org/10.1177/1536867X20909691

Czubala, W., Shepherd, B., & Wilson, J. S. (2009). Help or Hindrance? The Impact of Harmonised Standards on African Exports†. *Journal of African Economies*, *18*(5), 711–744. https://doi.org/10.1093/jae/ejp003

Danso-Abbeam, G., Ogundeji, A. A., & Fosu, S. (2022). Cashew contract farming in Ghana: Implications on farm performance and household welfare. *Journal of Agribusiness in Developing and Emerging Economies*, *ahead-of-print*(ahead-of-print). https://doi.org/10.1108/JADEE-12-2021-0339

Davidson, R., & MacKinnon, J. G. (2007). Moments of IV and JIVE estimators. *The Econometrics Journal*, *10*(3), 541–553.

Dawson, P. J. (2005). Agricultural exports and economic growth in less developed countries. *Agricultural Economics*, *33*(2), 145–152. https://doi.org/10.1111/j.1574-0862.2005.00358.x

de Brauw, A., & Bulte, E. (2021). *African Farmers, Value Chains and Agricultural Development: An Economic and Institutional Perspective*. Springer International Publishing. https://doi.org/10.1007/978-3-030-88693-6

Dedehouanou, S. F. A., Swinnen, J., & Maertens, M. (2013). Does Contracting Make Farmers Happy? Evidence from Senegal. *Review of Income and Wealth*, *59*(S1), S138–S160. https://doi.org/10.1111/roiw.12041

Dercon, S., Gilligan, D. O., Hoddinott, J., & Woldehanna, T. (2009). The Impact of Agricultural Extension and Roads on Poverty and Consumption Growth in Fifteen Ethiopian Villages. *American Journal of Agricultural Economics*, *91*(4), 1007–1021. https://doi.org/10.1111/j.1467-8276.2009.01325.x

Di Falco, S., Veronesi, M., & Yesuf, M. (2011). Does Adaptation to Climate Change Provide Food Security? A Micro-Perspective from Ethiopia. *American Journal of Agricultural Economics*, *93*(3), 829–846. https://doi.org/10.1093/ajae/aar006

Diao, X., McMillan, M., & Rodrik, D. (2019). The Recent Growth Boom in Developing Economies: A Structural-Change Perspective. In M. Nissanke & J. A. Ocampo (Eds.), *The Palgrave Handbook of Development Economics: Critical Reflections on Globalisation and Development* (pp. 281–334). Springer International Publishing. https://doi.org/10.1007/978-3-030-14000-7_9

Diao, X., McMillan, M., & Wangwe, S. (2018). Agricultural Labour Productivity and Industrialisation: Lessons for Africa. *Journal of African Economies*, *27*(1), 28–65. https://doi.org/10.1093/jae/ejx034

Dorward, A. (2001). The Effects of Transaction Costs, Power and Risk on Contractual Arrangements: A Conceptual Framework for Quantitative Analysis. *Journal of Agricultural Economics*, *52*(2), 59–73. https://doi.org/10.1111/j.1477-9552.2001.tb00925.x

Dubbert, C. (2019). Participation in contract farming and farm performance: Insights from cashew farmers in Ghana. *Agricultural Economics*, *50*(6), 749–763. https://doi.org/10.1111/agec.12522

Dubbert, C., & Abdulai, A. (2021). Does the Contract Type Matter? Impact of Marketing and Production Contracts on Cashew Farmers' Farm Performance in Ghana. *Journal of Agricultural & Food Industrial Organization*. https://doi.org/10.1515/jafio-2020-0040

Duval, Y., & Mengjing, K. (2017). *Digital Trade Facilitation: Paperless Trade in Regional Trade Agreements* (Working Paper No. 747; ADBI Working Paper Series, p. 43). Asian Development Bank Institute. https://www.econstor.eu/bitstream/10419/179203/1/adbi-wp747.pdf

Duval, Y., Utoktham, C., & Kravchenko, A. (2018). *Impact of implementation of digital trade facilitation on trade costs* (Working Paper No. 174; ARTNeT Working Paper Series, p. 58). UNESCAP. https://www.econstor.eu/bitstream/10419/177771/1/AWP174.pdf

Duval, Y., Wang, T., Utoktham, C., & Kravchenko, A. (2019). Next-generation Trade Facilitation for Asian Integration: Cross-border Paperless Trade. *Journal of Asian Economic Integration*, *1*(1), 11–31. https://doi.org/10.1177/2631684618821473

Ehrich, M., & Mangelsdorf, A. (2018). The Role of Private Standards for Manufactured Food Exports from Developing Countries. *World Development*, *101*, 16–27. https://doi.org/10.1016/j.worlddev.2017.08.004

Eichengreen, B., & Irwin, D. A. (1998). The role of history in bilateral trade flows. In J. Frankel (Ed.), *The Regionalization of the World Economy*. University of Chicago Press.

Ejones, F., Agbola, F. W., & Mahmood, A. (2021). Do regional trade agreements promote international trade? New empirical evidence from the East African Community. *The Journal of International Trade & Economic Development*, *0*(0), 1–34. https://doi.org/10.1080/09638199.2021.1930110

Engemann, H., & Jafari, Y. (2022). COVID-19 and changes in global agri-food trade. *Q Open*, *2*(1), qoac013. https://doi.org/10.1093/qopen/qoac013

Engman, M. (2009). The Economic Impact of Trade Facilitation. In *Overcoming Border Bottlenecks: The Costs and Benefits of Trade Facilitation* (pp. 81–112). OECD. https://doi.org/10.1787/9789264056954-4-en

Estmann, C., Sørensen, B. B., Ndulu, B., & Rand, J. (2022). Merchandise export diversification strategy for Tanzania: Promoting inclusive growth, economic complexity and structural change. *The World Economy*, *00*, 1–47. https://doi.org/10.1111/twec.13255

Fafchamps, M., & Minten, B. (2001). Social Capital and Agricultural Trade. *American Journal of Agricultural Economics*, *83*(3), 680–685. https://doi.org/10.1111/0002-9092.00190

Fally, T. (2015). Structural gravity and fixed effects. *Journal of International Economics*, *97*(1), 76–85. https://doi.org/10.1016/j.jinteco.2015.05.005

FAO. (2015). *Analysis of price incentives for cashew nuts in the United Republic of Tanzania 2005–2013* (Tecnical Note Series, MAFAP, p. 53). Food and Agriculture Organization of the United Nations (FAO). https://www.fao.org/3/i5041e/i5041e.pdf

FAO, IFAD, UNICEF, WFP, & WHO. (2020). *The State of Food Security and Nutrition in the World 2020: Transforming Food Systems for Affordable Healthy Diets* [Flagship Report].

Food and Agriculture Organization of the United Nations (FAO). https://doi.org/10.4060/ca9692en

FAO, IFAD, WFP, UNICEF, & WHO. (2022). *The State of Food Security and Nutrition in the World 2022: Repurposing Food and Agricultural Policies to Make Healthy Diets More Affordable* (p. 260). Food and Agriculture Organization of the United Nations (FAO). https://doi.org/10.4060/cc0639en

Feenstra, R. C. (2015). *Advanced International Trade: Theory and Evidence* (Second Edition). Princeton University Press.

Feyaerts, H., Broeck, G. V. den, & Maertens, M. (2020). Global and local food value chains in Africa: A review. *Agricultural Economics*, *51*(1), 143–157. https://doi.org/10.1111/agec.12546

Fiankor, D.-D. D., Curzi, D., & Olper, A. (2021). Trade, price and quality upgrading effects of agri-food standards. *European Review of Agricultural Economics*, *48*(4), 835–877. https://doi.org/10.1093/erae/jbaa026

Fidrmuc, J. (2004). The core and periphery of the world economy. *The Journal of International Trade & Economic Development*, *13*(1), 89–106. https://doi.org/10.1080/0963819042000213552

Fischer, E., & Qaim, M. (2012a). Linking Smallholders to Markets: Determinants and Impacts of Farmer Collective Action in Kenya. *World Development*, *40*(6), 1255–1268. https://doi.org/10.1016/j.worlddev.2011.11.018

Fischer, E., & Qaim, M. (2012b). Gender, agricultural commercialization, and collective action in Kenya. *Food Security*, *4*(3), 441–453. https://doi.org/10.1007/s12571-012-0199-7

Flach, L., & Unger, F. (2022). Quality and gravity in international trade. *Journal of International Economics*, *137*, 103578. https://doi.org/10.1016/j.jinteco.2022.103578

Fofack, H., Dzene, R., & Hussein, O. A. M. (2021). Estimating the Effect of AfCFTA on Intra-African Trade using Augmented GE-PPML. *Journal of African Trade*, *8*(2 (Special Issue)), 62–76. https://doi.org/10.2991/jat.k.211122.001

Frankel, J. A., & Romer, D. H. (1999). Does Trade Cause Growth? *American Economic Review*, *89*(3), 379–399. https://doi.org/10.1257/aer.89.3.379

Frankel, J., Stein, E., & Wei, S. (1995). Trading blocs and the Americas: The natural, the unnatural, and the super-natural. *Journal of Development Economics*, *47*(1), 61–95. https://doi.org/10.1016/0304-3878(95)00005-4

Frelat, R., Lopez-Ridaura, S., Giller, K. E., Herrero, M., Douxchamps, S., Djurfeldt, A. A., Erenstein, O., Henderson, B., Kassie, M., Paul, B. K., Rigolot, C., Ritzema, R. S., Rodriguez, D., van Asten, P. J. A., & van Wijk, M. T. (2016). Drivers of household food availability in sub-Saharan Africa based on big data from small farms. *Proceedings of the National Academy of Sciences*, *113*(2), 458–463. https://doi.org/10.1073/pnas.1518384112

Freund, C. L., & Weinhold, D. (2004). The effect of the Internet on international trade. *Journal of International Economics*, *62*(1), 171–189. https://doi.org/10.1016/S0022-1996(03)00059-X

Fukase, E., & Martin, W. (2020). Economic growth, convergence, and world food demand and supply. *World Development*, *132*, 104954. https://doi.org/10.1016/j.worlddev.2020.104954

Gaulier, G., & Zignago, S. (2010). *BACI: International Trade Database at the Product-Level. The 1994-2007 Version* (Working Paper 2010=23; CEPII Working Paper). CEPII. http://www.cepii.fr/PDF_PUB/wp/2010/wp2010-23.pdf

Gelaw, F., Speelman, S., & Van Huylenbroeck, G. (2016). Farmers' marketing preferences in local coffee markets: Evidence from a choice experiment in Ethiopia. *Food Policy*, *61*, 92–102. https://doi.org/10.1016/j.foodpol.2016.02.006

Gereffi, G., Humphrey, J., Kaplinsky, R., & Sturgeon*, T. J. (2001). Introduction: Globalisation, Value Chains and Development. *IDS Bulletin*, *32*(3), 1–8. https://doi.org/10.1111/j.1759-5436.2001.mp32003001.x

Gereffi, G., Humphrey, J., & Sturgeon, T. (2005). The governance of global value chains. *Review of International Political Economy*, *12*(1), 78–104. https://doi.org/10.1080/09692290500049805

Gibson, J., & McKenzie, D. (2014). The development impact of a best practice seasonal worker policy. *The Review of Economics and Statistics*, *96*(2), 229–243.

Glaeser, E. L. (2010). *Agglomeration economics*. The University of Chicago Press.

Goedhuys, M., & Sleuwaegen, L. (2016). International standards certification, institutional voids and exports from developing country firms. *International Business Review*, *25*(6), 1344–1355. https://doi.org/10.1016/j.ibusrev.2016.04.006

Gollin, D., Lagakos, D., & Waugh, M. E. (2014a). Agricultural Productivity Differences across Countries. *American Economic Review*, *104*(5), 165–170. https://doi.org/10.1257/aer.104.5.165

Gollin, D., Lagakos, D., & Waugh, M. E. (2014b). The Agricultural Productivity Gap*. *The Quarterly Journal of Economics*, *129*(2), 939–993. https://doi.org/10.1093/qje/qjt056

Gollin, D., Parente, S., & Rogerson, R. (2002). The Role of Agriculture in Development. *The American Economic Review*, *92*(2), 160–164.

Gollin, D., & Rogerson, R. (2014). Productivity, transport costs and subsistence agriculture. *Journal of Development Economics*, *107*, 38–48. https://doi.org/10.1016/j.jdeveco.2013.10.007

Gollin, D., & Udry, C. (2021). Heterogeneity, Measurement Error, and Misallocation: Evidence from African Agriculture. *Journal of Political Economy*, *129*(1), 1–80. https://doi.org/10.1086/711369

Gouel, C., & Guimbard, H. (2019). Nutrition Transition and the Structure of Global Food Demand. *American Journal of Agricultural Economics*, *101*(2), 383–403. https://doi.org/10.1093/ajae/aay030

Cashewnut Industry Act, no. No. 18, Parliament of Tanzania (2009). https://trade.tanzania.go.tz/media/The%20Cashewnut%20Industry%20Act,%202009.pdf

Government of Tanzania. (2015). *Agricultural Sector Development Strategy II (2015/2016—2024/2025)* [Government Programme Document]. Ministry of Agriculture. http://extwprlegs1.fao.org/docs/pdf/tan160643.pdf

Government of Tanzania. (2021). *National Five Year Development Plan 2021/22—2025/26: Realising Competitiveness and Industrialisation for Human Development*. Ministry of Finance and Planning. https://mof.go.tz/docs/news/FYDP%20III%20English.pdf

Gözgör, G., & Can, M. (2017). Causal Linkages among the Product Diversification of Exports, Economic Globalization and Economic Growth. *Review of Development Economics*, *21*(3), 888–908. https://doi.org/10.1111/rode.12301

Graeub, B. E., Chappell, M. J., Wittman, H., Ledermann, S., Kerr, R. B., & Gemmill-Herren, B. (2016). The State of Family Farms in the World. *World Development*, *87*, 1–15. https://doi.org/10.1016/j.worlddev.2015.05.012

Grainger, A. (2011). Trade Facilitation: A Conceptual Review. *Journal of World Trade*, *45*(1), 39–62.

Grant, J. H., & Lambert, D. M. (2008). Do Regional Trade Agreements Increase Members' Agricultural Trade? *American Journal of Agricultural Economics*, *90*(3), 765–782.

Grosh, B. (1994). Contract Farming in Africa: An Application of the New Institutional Economics. *Journal of African Economies*, *3*(2), 231–261. https://doi.org/10.1093/oxfordjournals.jae.a036805

Haq, Z. U., Meilke, K., & Cranfield, J. (2013). Selection bias in a gravity model of agrifood trade. *European Review of Agricultural Economics*, *40*(2), 331–360. https://doi.org/10.1093/erae/jbs028

Haveman, J., & Hummels, D. (2004). Alternative Hypotheses and the Volume of Trade: The Gravity Equation and the Extent of Specialization. *The Canadian Journal of Economics / Revue Canadienne d'Economique*, *37*(1), 199–218.

Heckman, J. J., Ichimura, H., & Todd, P. E. (1997). Matching as an Econometric Evaluation Estimator: Evidence from Evaluating a Job Training Programme. *The Review of Economic Studies*, *64*(4), 605–654. JSTOR. https://doi.org/10.2307/2971733

Henderson, J. V., Squires, T., Storeygard, A., & Weil, D. (2018). The Global Distribution of Economic Activity: Nature, History, and the Role of Trade1. *The Quarterly Journal of Economics*, *133*(1), 357–406. https://doi.org/10.1093/qje/qjx030

Herrmann, R., Nkonya, E., & Faße, A. (2018). Food value chain linkages and household food security in Tanzania. *Food Security*, *10*(4), 827–839. https://doi.org/10.1007/s12571-018-0792-5

Hirvonen, K., Bai, Y., Headey, D., & Masters, W. A. (2020). Affordability of the EAT–Lancet reference diet: A global analysis. *The Lancet Global Health*, *8*(1), e59–e66. https://doi.org/10.1016/S2214-109X(19)30447-4

Hollinger, F., & Staatz, J. M. (Eds.). (2015). *Agricultural growth in West Africa: Market and policy drivers*. African Development Bank and the Food and Agriculture Organization of the United Nations. https://www.fao.org/3/i4337e/i4337e.pdf

Huellen, S. van, & Abubakar, F. M. (2021). Potential for Upgrading in Financialised Agri-food Chains: The Case of Ghanaian Cocoa. *The European Journal of Development Research*, *33*(2). https://doi.org/10.1057/s41287-020-00351-3

Ibrahim, G., Simbanegavi, W., Prakash, A., Davis, W., Wasike, W., & Patel, A. (2019). *Industrial Development and ICT in Africa: Opportunities, Challenges and Way Forward* (p. 15). JICA. https://www.jica.go.jp/jica-ri/publication/booksandreports/l75nbg000017wb6r-att/TF5_web_0603_0006.pdf

Ichino, A., Mealli, F., & Nannicini, T. (2008). From temporary help jobs to permanent employment: What can we learn from matching estimators and their sensitivity? *Journal of Applied Econometrics*, *23*(3), 305–327. https://doi.org/10.1002/jae.998

Imbens, G. W. (2006). Nonparametric Estimation of Average Treatment Effects Under Exogeneity: A Review. *The Review of Economics and Statistics*, *86*(1), 4–29. http://dx.doi.org/10.1162/003465304323023651

Imbens, G. W., & Wooldridge, J. M. (2009). Recent Developments in the Econometrics of Program Evaluation. *Journal of Economic Literature*, *47*(1), 5–86. https://doi.org/10.1257/jel.47.1.5

International Finance Corporation. (2013). *Working with Smallholders: A Handbook for Firms Building Sustainable Supply Chains*. The World Bank Group. 10.1596/978-1-4648 -1277-4

Janssens, C., Havlík, P., Krisztin, T., Baker, J., Frank, S., Hasegawa, T., Leclère, D., Ohrel, S., Ragnauth, S., Schmid, E., Valin, H., Van Lipzig, N., & Maertens, M. (2020). Global hunger and climate change adaptation through international trade. *Nature Climate Change*, *10*(9), 829–835. https://doi.org/10.1038/s41558-020-0847-4

Janvry, A. de, Fafchamps, M., & Sadoulet, E. (1991). Peasant Household Behaviour with Missing Markets: Some Paradoxes Explained. *The Economic Journal*, *101*(409), 1400–1417. https://doi.org/10.2307/2234892

Jensen, R. (2007). The Digital Provide: Information (Technology), Market Performance, and Welfare in the South Indian Fisheries Sector. *The Quarterly Journal of Economics*, *122*(3), 879–924. https://doi.org/10.1162/qjec.122.3.879

Jensen, R. (2010). Information, efficiency, and welfare in agricultural markets. *Agricultural Economics*, *41*(s1), 203–216. https://doi.org/10.1111/j.1574-0862.2010.00501.x

Johnston, B. F., & Mellor, J. W. (1961). The Role of Agriculture in Economic Development. *The American Economic Review*, *51*(4), 566–593.

Kabasa, J. D., Kirsten, J., & Minde, I. (2015). Implications of changing agri-food system structure for agricultural education and training in Sub-Saharan Africa. *Journal of Agribusiness in Developing and Emerging Economies*, *5*(2), 190–199. https://doi.org/10.1108/JADEE-03-2015-0016

Kabir, M., Salim, R., & Al-Mawali, N. (2017). The gravity model and trade flows: Recent developments in econometric modeling and empirical evidence. *Economic Analysis and Policy*, *56*, 60–71. https://doi.org/10.1016/j.eap.2017.08.005

Kaplinsky, R., & Morris, M. (2000). *A handbook for value chain research* (Vol. 113). Institute of Development Studies: University of Sussex.

Kareem, F. O., Martínez-Zarzoso, I., & Brümmer, B. (2018). Protecting health or protecting imports? Evidence from EU non-tariff measures. *International Review of Economics & Finance*, *53*, 185–202. https://doi.org/10.1016/j.iref.2017.08.012

Kennedy, G., Ballard, T., & Dop, M. C. (2010). *Guidelines for measuring household and individual dietary diversity*. Food and Agriculture Organization of the United Nations.

Key, N., & Runsten, D. (1999). Contract Farming, Smallholders, and Rural Development in Latin America: The Organization of Agroprocessing Firms and the Scale of Outgrower Production. *World Development*, *27*(2), 381–401. https://doi.org/10.1016/S0305-750X(98)00144-2

Key, N., Sadoulet, E., & Janvry, A. D. (2000). Transactions Costs and Agricultural Household Supply Response. *American Journal of Agricultural Economics*, *82*(2), 245–259. https://doi.org/10.1111/0002-9092.00022

Kilama, B. (2013). *The diverging South: Comparing the cashew sectors of Tanzania and Vietnam*. Leiden University.

Kirsten, J., & Sartorius, K. (2002). Linking agribusiness and small-scale farmers in developing countries: Is there a new role for contract farming? *Development Southern Africa*, *19*(4), 503–529. https://doi.org/10.1080/0376835022000019428

Kirui, O. K., Kornher, L., & Bekchanov, M. (2023). Productivity growth and the role of mechanisation in African agriculture. *Agrekon*, *0*(0), 1–18. https://doi.org/10.1080/03031853.2023.2176894

Klasen, S., Martínez-Zarzoso, I., Nowak-Lehmann, F., & Bruckner, M. (2021). Does the designation of least developed country status promote exports? *The Journal of International Trade & Economic Development*, *30*(2), 157–177. https://doi.org/10.1080/09638199.2020.1831042

Kornher, L., & von Braun, J. (2020). *EU Common Agricultural Policy—Impacts on Trade with Africa and African Agricultural Development* (No. 294; ZEF Discussion Papers on Development Policy No. 204, p. 65). University of Bonn: ZEF.

Kubik, Z., Getahun, T., Omari, R., & Zlaoui, M. O. (2022). *Can the agroprocessing sector create jobs in Africa? Evidence from Ethiopia, Ghana and Tunisia* (Working Paper WP 215; ZEF Working Paper Series). Center for Development Research (ZEF), University of Bonn. https://www.zef.de/fileadmin/user_upload/ZEF_Working_Paper_215.pdf

Kumar Roy, C., & Xiaoling, H. (2020). Effects of Paperless Trade Policy and Aid for Trade on Export Performance: Evidence from SASEC And CAREC Countries. *Asian Development Policy Review*, *8*(1), 61–74. https://doi.org/10.18488/journal.107.2020.81.61.74

Kuzilwa, J. A., Fold, N., Henningsen, A., & Larsen, M. N. (Eds.). (2019). *Contract Farming and the Development of Smallholder Agricultural Businesses: Improving markets and value chains in Tanzania*. Routledge. https://doi.org/10.4324/9781315651422

Lagakos, D., & Waugh, M. E. (2013). Selection, Agriculture, and Cross-Country Productivity Differences. *American Economic Review*, *103*(2), 948–980. https://doi.org/10.1257/aer.103.2.948

Lambert, D., & McKoy, S. (2009). Trade Creation and Diversion Effects of Preferential Trade Associations on Agricultural and Food Trade. *Journal of Agricultural Economics*, *60*(1), 17–39. https://doi.org/10.1111/j.1477-9552.2008.00184.x

Lancaster, K. J. (1966). A New Approach to Consumer Theory. *Journal of Political Economy*, *74*(2), 132–157. JSTOR.

Laryea, E. (2005). Facilitating paperless international trade: A survey of Law and Policy in Asia. *International Review of Law, Computers & Technology*, *19*(2), 121–142. https://doi.org/10.1080/13600860500131200

Lewis, G. (2009). *The Impact of ICT on Customs*. *3*(1), 3–11.

Lewis, W. A. (1954). Economic Development with Unlimited Supplies of Labour. *The Manchester School*, *22*(2), 139–191. https://doi.org/10.1111/j.1467-9957.1954.tb00021.x

Li, Y., & Beghin, J. C. (2012). A meta-analysis of estimates of the impact of technical barriers to trade. *Journal of Policy Modeling*, *34*(3), 497–511. https://doi.org/10.1016/j.jpolmod.2011.11.001

Lin, B., Wang, X., Jin, S., Yang, W., & Li, H. (2022). Impacts of cooperative membership on rice productivity: Evidence from China. *World Development*, *150*, 105669. https://doi.org/10.1016/j.worlddev.2021.105669

Lin, F. (2015). Estimating the effect of the Internet on international trade. *The Journal of International Trade & Economic Development*, *24*(3), 409–428. https://doi.org/10.1080/09638199.2014.881906

Liverpool-Tasie, L. S. O., Reardon, T., & Belton, B. (2021). "Essential non-essentials": COVID-19 policy missteps in Nigeria rooted in persistent myths about African food supply chains. *Applied Economic Perspectives and Policy*, *43*(1), 205–224. https://doi.org/10.1002/aepp.13139

Lowder, S. K., Sánchez, M. V., & Bertini, R. (2021). Which farms feed the world and has farmland become more concentrated? *World Development*, *142*, 105455. https://doi.org/10.1016/j.worlddev.2021.105455

Lowder, S. K., Skoet, J., & Raney, T. (2016). The Number, Size, and Distribution of Farms, Smallholder Farms, and Family Farms Worldwide. *World Development*, *87*, 16–29. https://doi.org/10.1016/j.worlddev.2015.10.041

Ma, A., & Van Assche, A. (2011). Global Production Networks in the Postcrisis Era. In M. Haddad & B. Shepherd, *Managing Openness: Trade and Outward-oriented Growth After the Crisis*. World Bank.

Ma, W., & Abdulai, A. (2017). The economic impacts of agricultural cooperatives on smallholder farmers in rural China. *Agribusiness*, *33*(4), 537–551. https://doi.org/10.1002/agr.21522

Macchiavello, R., & Morjaria, A. (2015). The Value of Relationships: Evidence from a Supply Shock to Kenyan Rose Exports. *American Economic Review*, *105*(9), 2911–2945. https://doi.org/10.1257/aer.20120141

Macchiavello, R., & Morjaria, A. (2021). Competition and Relational Contracts in the Rwanda Coffee Chain. *The Quarterly Journal of Economics*, *136*(2), 1089–1143. https://doi.org/10.1093/qje/qjaa048

Maertens, A., & Barrett, C. B. (2013). Measuring Social Networks' Effects on Agricultural Technology Adoption. *American Journal of Agricultural Economics*, *95*(2), 353–359. https://doi.org/10.1093/ajae/aas049

Maertens, M., & Swinnen, J. F. M. (2009). Trade, Standards, and Poverty: Evidence from Senegal. *World Development*, *37*(1), 161–178. https://doi.org/10.1016/j.worlddev.2008.04.006

Maertens, M., & Vande Velde, K. (2017). Contract-farming in Staple Food Chains: The Case of Rice in Benin. *World Development*, *95*, 73–87. https://doi.org/10.1016/j.worlddev.2017.02.011

Manh Hung, N., & Makdissi, P. (2004). Escaping the poverty trap in a developing rural economy. *Canadian Journal of Economics*, *37*(1), 123–139. https://doi.org/10.1111/j.0008-4085.2004.006_1.x

Markelova, H., Meinzen-Dick, R., Hellin, J., & Dohrn, S. (2009). Collective action for smallholder market access. *Food Policy*, *34*(1), 1–7. https://doi.org/10.1016/j.foodpol.2008.10.001

Masters, W. A., Rosenblum, N. Z., & Alemu, R. G. (2018). Agricultural Transformation, Nutrition Transition and Food Policy in Africa: Preston Curves Reveal New Stylised Facts. *The Journal of Development Studies*, *54*(5), 788–802. https://doi.org/10.1080/00220388.2018.1430768

McMillan, M., Page, J., & Wangwe, S. (2017). Unlocking Tanzania's Manufacturing Potential. In C. S. Adam, P. Collier, & B. Ndulu, *Tanzania: The Path to Prosperity* (pp. 151–169). Oxford University Press.

McMillan, M., Rodrik, D., & Verduzco-Gallo, Í. (2014). Globalization, Structural Change, and Productivity Growth, with an Update on Africa. *World Development*, *63*, 11–32. https://doi.org/10.1016/j.worlddev.2013.10.012

McMillan, M., Rodrik, D., & Welch, K. H. (2002). *When Economic Reform Goes Wrong: Cashews in Mozambique* (No. 9117; NBER Working Paper Series). National Bureau of Economic Research. https://www.nber.org/system/files/working_papers/w9117/w9117.pdf

McMillan, M., & Zeufack, A. (2022). Labor Productivity Growth and Industrialization in Africa. *Journal of Economic Perspectives*, *36*(1), 3–32. https://doi.org/10.1257/jep.36.1.3

Meemken, E.-M., & Bellemare, M. F. (2020). Smallholder farmers and contract farming in developing countries. *Proceedings of the National Academy of Sciences*, *117*(1), 259–264. https://doi.org/10.1073/pnas.1909501116

Mellor, J. W. (2017). *Agricultural Development and Economic Transformation: Promoting Growth with Poverty Reduction*. Palgrave Macmillan. https://doi.org/10.1007/978-3-319-65259-7

Melo, J. de, & Nicita, A. (2018). *Non-Tariff Measures: Data and Quantitative Tools of Analysis* (Working Papers P218; Working Paper Development Policies). FERDI.

https://ferdi.fr/dl/df-pwFsQKwSBf2Djde77ELncfwt/ferdi-p218-non-tariff-measures-data-and-quantitative-tools-of-analysis.pdf

Melo, J. de, & Shepherd, B. (2018). The Economics of Non-Tariff Measures: A Primer. In *Non-Tariff Measures: Economic Assessment and Policy Options for Development* (pp. 121–157). UNCTAD. https://unctad.org/system/files/official-document/ditctab2018d3_en.pdf

Mergenthaler, M., Weinberger, K., & Qaim, M. (2009). The food system transformation in developing countries: A disaggregate demand analysis for fruits and vegetables in Vietnam. *Food Policy*, *34*(5), 426–436. https://doi.org/10.1016/j.foodpol.2009.03.009

Miassi, Y. E., Dossa, F. K., Labiyi, I. A., Dossouhoui, S., & Yabi, J. A. (2019). Contract Farming and Cashew Production in North Benin: Socio-economic importance and determinants of profitability. *Journal of Agricultural Research Advances*. https://hal.archives-ouvertes.fr/hal-02095883

Miassi, Y. E. S., & Dossa, F. K. (2019). Socio-Economic Determinants of the Adoption of Agricultural Contracts: Case of Cashew Farmers in North-Eastern Benin. *Journal of Agricultural Research Advances*, *1*(1), 29–38.

Michelson, H. C. (2017). Influence of Neighbor Experience and Exit on Small Farmer Market Participation. *American Journal of Agricultural Economics*, *99*(4), 952–970. https://doi.org/10.1093/ajae/aaw097

Michler, J. D., & Wu, S. Y. (2020). Relational Contracts in Agriculture: Theory and Evidence. *Annual Review of Resource Economics*, *12*(1), 111–127. https://doi.org/10.1146/annurev-resource-101719-034514

Milner, C., Morrissey, O., & Rudaheranwa, N. (2000). Policy and Non-Policy Barriers to Trade and Implicit Taxation of Exports in Uganda. *Journal of Development Studies*, *37*(2), 67–90. https://doi.org/10.1080/713600069

Mishra, A. K., Kumar, A., Joshi, P. K., & D'souza, A. (2016). Impact of contracts in high yielding varieties seed production on profits and yield: The case of Nepal. *Food Policy*, *62*, 110–121. https://doi.org/10.1016/j.foodpol.2016.05.005

Mishra, A. K., Kumar, A., Joshi, P. K., D'Souza, A., & Tripathi, G. (2018). How can organic rice be a boon to smallholders? Evidence from contract farming in India. *Food Policy*, *75*, 147–157. https://doi.org/10.1016/j.foodpol.2018.01.007

Miyata, S., Minot, N., & Hu, D. (2009). Impact of Contract Farming on Income: Linking Small Farmers, Packers, and Supermarkets in China. *World Development*, *37*(11), 1781–1790. https://doi.org/10.1016/j.worlddev.2008.08.025

Mmbando, F. E., Wale, E. Z., & Baiyegunhi, L. J. S. (2015). Welfare impacts of smallholder farmers' participation in maize and pigeonpea markets in Tanzania. *Food Security*, *7*(6), 1211–1224. https://doi.org/10.1007/s12571-015-0519-9

Moenius, J. (2004). *Information Versus Product Adaptation: The Role of Standards in Trade* (SSRN Scholarly Paper ID 608022). Social Science Research Network. https://doi.org/10.2139/ssrn.608022

Montalbano, P., Pietrelli, R., & Salvatici, L. (2018). Participation in the market chain and food security: The case of the Ugandan maize farmers. *Food Policy*, *76*, 81–98. https://doi.org/10.1016/j.foodpol.2018.03.008

Morrissy, J. D. (1974). *Agricultural Modernization Through Production Contracting: Role of the Fruit and Vegetable Processor in Mexico and Central America*. Praeger Publishers Inc.

Msami, J., & Wangwe, S. (2016). Industrial Development in Tanzania. In C. Newman, J. Page, J. Rand, A. Shimeles, M. Söderbom, & F. Tarp, *Manufacturing Transformation: Comparative Studies of Industrial Development in Africa and Emerging Asia*. Oxford University Press.

https://oxford.universitypressscholarship.com/view/10.1093/acprof:oso/9780198776987.001.0001/acprof-9780198776987

Mujahid, I., & Kalkuhl, M. (2016). Do Trade Agreements Increase Food Trade? *The World Economy*, *39*(11), 1812–1833. https://doi.org/10.1111/twec.12324

Muriithi, B., & Kabubo-Mariara, J. (2022). The Dynamics and Role of Gender in High-Value Avocado Farming in Kenya. *The European Journal of Development Research*, *34*(5), 2272–2304. https://doi.org/10.1057/s41287-021-00484-z

Mutabazi, K. D., Sieber, S., Maeda, C., & Tscherning, K. (2015). Assessing the determinants of poverty and vulnerability of smallholder farmers in a changing climate: The case of Morogoro region, Tanzania. *Regional Environmental Change*, *15*(7), 1243–1258. https://doi.org/10.1007/s10113-015-0772-7

Mwambi, M., Oduol, J., Mshenga, P., & Mwanarusi, S. (2016). Does Contract Farming Improve Smallholder Farmers Income? The Case of Avocado Farming in Kenya. *Journal of Agribusiness in Developing and Emerging Economies*, *6*(1), 2–20.

Namonje-Kapembwa, T., Black, R., & TS, J. (2017). Does Late Delivery of Subsidized Fertilizer Affect Smallholder Maize Productivity and Production. *Agricultural Research & Technology: Open Access Journal*, *11*(2). https://doi.org/10.19080/ARTOAJ.2017.11.555810

Narayanan, S. (2014). Profits from participation in high value agriculture: Evidence of heterogeneous benefits in contract farming schemes in Southern India. *Food Policy*, *44*, 142–157. https://doi.org/10.1016/j.foodpol.2013.10.010

NBS. (2021). *National Sample Census of Agriculture 2019/20: National Report* (p. 931). National Bureau of Statistics (NBS). https://www.nbs.go.tz/nbs/takwimu/Agriculture/2019-20_Agri_Census_Main_Report.pdf

Ngatunga, E. L., Dondeyne, S., & Deckers, J. A. (2003). Is sulphur acidifying cashew soils of South Eastern Tanzania? *Agriculture, Ecosystems & Environment*, *95*(1), 179–184. https://doi.org/10.1016/S0167-8809(02)00104-4

Nguyen, A. T., Dzator, J., & Nadolny, A. (2015). Does contract farming improve productivity and income of farmers?: A review of theory and evidence. *The Journal of Developing Areas*, *49*(6), 531–538. https://doi.org/10.1353/jda.2015.0094

Niehans, J. (1987). Transaction Costs. In *The New Palgrave Dictionary of Economics* (pp. 1–6). Palgrave Macmillan UK. https://doi.org/10.1057/978-1-349-95121-5_1682-1

Niu, Z., Liu, C., Gunessee, S., & Milner, C. (2018). Non-tariff and overall protection: Evidence across countries and over time. *Review of World Economics*, *154*(4), 675–703. https://doi.org/10.1007/s10290-018-0317-5

Nooteboom, B. (1993). Firm size effects on transaction costs. *Small Business Economics*, *5*(4), 283–295. https://doi.org/10.1007/BF01516249

Ochieng, D. O., Veettil, P. C., & Qaim, M. (2017). Farmers' preferences for supermarket contracts in Kenya. *Food Policy*, *68*, 100–111. https://doi.org/10.1016/j.foodpol.2017.01.008

Ogutu, S. O., Ochieng, D. O., & Qaim, M. (2020). Supermarket contracts and smallholder farmers: Implications for income and multidimensional poverty. *Food Policy*, 101940. https://doi.org/10.1016/j.foodpol.2020.101940

Ogutu, S. O., & Qaim, M. (2019). Commercialization of the small farm sector and multidimensional poverty. *World Development*, *114*, 281–293. https://doi.org/10.1016/j.worlddev.2018.10.012

Ola, O., & Menapace, L. (2020a). Revisiting constraints to smallholder participation in high-value markets: A best-worst scaling approach. *Agricultural Economics*, *51*(4), 595–608. https://doi.org/10.1111/agec.12574

Ola, O., & Menapace, L. (2020b). Smallholders' perceptions and preferences for market attributes promoting sustained participation in modern agricultural value chains. *Food Policy*, *97*, 101962. https://doi.org/10.1016/j.foodpol.2020.101962

Otsuka, K., Nakano, Y., & Takahashi, K. (2016). Contract Farming in Developed and Developing Countries. *Annual Review of Resource Economics*, *8*(1), 353–376. https://doi.org/10.1146/annurev-resource-100815-095459

Otsuki, T., Wilson, J. S., & Sewadeh, M. (2001). Saving two in a billion: Quantifying the trade effect of European food safety standards on African exports. *Food Policy*, *26*(5), 495–514. https://doi.org/10.1016/S0306-9192(01)00018-5

Oya, C. (2012). Contract Farming in Sub-Saharan Africa: A Survey of Approaches, Debates and Issues. *Journal of Agrarian Change*, *12*(1), 1–33. https://doi.org/10.1111/j.1471-0366.2011.00337.x

Pernechele, V., Fontes, F., Baborska, R., Nkuingoua Nana, J. C., Pan, X., & Tuyishime, C. (2021). *Public expenditure on food and agriculture in sub-Saharan Africa: Trends, Challenges and Priorities*. Food and Agriculture Organization of the United Nations (FAO). http://www.fao.org/documents/card/en/c/cb4492en

Pettersson, J., & Johansson, L. (2013). Aid, Aid for Trade, and bilateral trade: An empirical study. *The Journal of International Trade & Economic Development*, *22*(6), 866–894. https://doi.org/10.1080/09638199.2011.613998

Philippidis, G., Resano-Ezcaray, H., & Sanjuán-López, A. I. (2013). Capturing zero-trade values in gravity equations of trade: An analysis of protectionism in agro-food sectors. *Agricultural Economics*, *44*(2), 141–159. https://doi.org/10.1111/agec.12000

Ponte, S., & Brockington, D. (2020). From pyramid to pointed egg? A 20-year perspective on poverty, prosperity, and rural transformation in Tanzania. *African Affairs*, *119*(475), 203–223. https://doi.org/10.1093/afraf/adaa002

Porteous, O. (2019). High Trade Costs and Their Consequences: An Estimated Dynamic Model of African Agricultural Storage and Trade. *American Economic Journal: Applied Economics*, *11*(4), 327–366. https://doi.org/10.1257/app.20170442

Porter, M. E. (1985). *The Competitive Advantage: Creating and Sustaining Superior Performance*. Free Press.

Porto, G., Chauvin, N., & Olarreaga, M. (2011). *Supply Chains in Export Agriculture, Competition, and Poverty in Sub-Saharan Africa*. The World Bank. https://academic.oup.com/erae/article-lookup/doi/10.1093/erae/jbr049

Putterman, L. (1995). Economic reform and smallholder agriculture in Tanzania: A discussion of recent market liberalization, road rehabilitation, and technology dissemination efforts. *World Development*, *23*(2), 311–326. https://doi.org/10.1016/0305-750X(94)00121-E

Rada, N. E., & Fuglie, K. O. (2019). New perspectives on farm size and productivity. *Food Policy*, *84*, 147–152. https://doi.org/10.1016/j.foodpol.2018.03.015

Ragasa, C., Lambrecht, I., & Kufoalor, D. S. (2018). Limitations of Contract Farming as a Pro-poor Strategy: The Case of Maize Outgrower Schemes in Upper West Ghana. *World Development*, *102*, 30–56. https://doi.org/10.1016/j.worlddev.2017.09.008

Ragasa, C., & Mazunda, J. (2018). The impact of agricultural extension services in the context of a heavily subsidized input system: The case of Malawi. *World Development*, *105*, 25–47. https://doi.org/10.1016/j.worlddev.2017.12.004

Rao, E. J. O., Brümmer, B., & Qaim, M. (2012). Farmer Participation in Supermarket Channels, Production Technology, and Efficiency: The Case of Vegetables in Kenya. *American Journal of Agricultural Economics*, *94*(4), 891–912. https://doi.org/10.1093/ajae/aas024

Rao, E. J. O., & Qaim, M. (2011). Supermarkets, Farm Household Income, and Poverty: Insights from Kenya. *World Development*, *39*(5), 784–796. https://doi.org/10.1016/j.worlddev.2010.09.005

Reardon, T., Barrett, C. B., Berdegué, J. A., & Swinnen, J. F. M. (2009). Agrifood Industry Transformation and Small Farmers in Developing Countries. *World Development*, *37*(11), 1717–1727. https://doi.org/10.1016/j.worlddev.2008.08.023

Reardon, T., Liverpool-Tasie, L. S. O., & Minten, B. (2021). Quiet Revolution by SMEs in the midstream of value chains in developing regions: Wholesale markets, wholesalers, logistics, and processing. *Food Security*, *13*(6), 1577–1594. https://doi.org/10.1007/s12571-021-01224-1

Reardon, T., & Timmer, C. P. (2012). The Economics of the Food System Revolution. *Annual Review of Resource Economics*, *4*(1), 225–264. https://doi.org/10.1146/annurev.resource.050708.144147

Reardon, T., Tschirley, D., Liverpool-Tasie, L. S. O., Awokuse, T., Fanzo, J., Minten, B., Vos, R., Dolislager, M., Sauer, C., Dhar, R., Vargas, C., Lartey, A., Raza, A., & Popkin, B. M. (2021). The processed food revolution in African food systems and the double burden of malnutrition. *Global Food Security*, *28*, 100466. https://doi.org/10.1016/j.gfs.2020.100466

Reil, S., Hamdani, K., Barreiros, L. E., Berger, A., Huerta, R. C., Kagan, Y., Sauvant, K. P., Steneri, P., & Zhao, Q. (2022). *What Foreign Investors Want: Findings From an Investor Survey of Investment Facilitation Measures in Latin America and the Caribbean*. International Trade Centre. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4018005

Rodrik, D. (2016). Premature deindustrialization. *Journal of Economic Growth*, *21*(1), 1–33. https://doi.org/10.1007/s10887-015-9122-3

Roessler, P., Pengl, Y. I., Marty, R., Titlow, K. S., & van de Walle, N. (2022). The cash crop revolution, colonialism and economic reorganization in Africa. *World Development*, *158*, 105934. https://doi.org/10.1016/j.worlddev.2022.105934

Rosenbaum, P. R. (2005). Sensitivity Analysis in Observational Studies. In B. S. Everitt & D. C. Howell, *Encyclopedia of Statistics in Behavioral Science* (pp. 1809–1814). John Wiley & Sons. https://doi.org/10.1002/0470013192.bsa606

Rosenbaum, P. R., & Rubin, D. B. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika*, *70*(1), 41–55. https://doi.org/10.1093/biomet/70.1.41

Rosenbaum, P. R., & Rubin, D. B. (1985). Constructing a Control Group Using Multivariate Matched Sampling Methods That Incorporate the Propensity Score. *The American Statistician*, *39*(1), 33–38. https://doi.org/10.2307/2683903

Ruml, A., & Qaim, M. (2020). Effects of marketing contracts and resource-providing contracts in the African small farm sector: Insights from oil palm production in Ghana. *World Development*, *136*, 105110. https://doi.org/10.1016/j.worlddev.2020.105110

Ruml, A., & Qaim, M. (2021a). New evidence regarding the effects of contract farming on agricultural labor use. *Agricultural Economics*, *52*(1), 51–66. https://doi.org/10.1111/agec.12606

Ruml, A., & Qaim, M. (2021b). Smallholder Farmers' Dissatisfaction with Contract Schemes in Spite of Economic Benefits: Issues of Mistrust and Lack of Transparency. *The Journal of Development Studies*, *57*(7), 1106–1119. https://doi.org/10.1080/00220388.2020.1850699

Ruml, A., Ragasa, C., & Qaim, M. (2022). Contract farming, contract design and smallholder livelihoods. *Australian Journal of Agricultural and Resource Economics*, *66*(1), 24–43. https://doi.org/10.1111/1467-8489.12462

Rutakumwa, R., Mugisha, J. O., Bernays, S., Kabunga, E., Tumwekwase, G., Mbonye, M., & Seeley, J. (2020). Conducting in-depth interviews with and without voice recorders: A comparative analysis. *Qualitative Research*, *20*(5), 565–581. https://doi.org/10.1177/1468794119884806

Saha, A., Castro, A., Carreras, M., & Guariso, D. (2020). *Trade, Technology and Absorptive Capacity: Firm-level evidence from Tanzanian Textiles Apparel* (Working Paper No. 2020/96; WIDER Working Paper Series). UNU-WIDER. https://doi.org/10.35188/UNU-WIDER/2020/853-5

Santeramo, F. G., & Lamonaca, E. (2019). The Effects of Non-tariff Measures on Agri-food Trade: A Review and Meta-analysis of Empirical Evidence. *Journal of Agricultural Economics*, *70*(3), 595–617. https://doi.org/10.1111/1477-9552.12316

Santos Silva, J. M. C., & Tenreyro, S. (2010). On the existence of the maximum likelihood estimates in Poisson regression. *Economics Letters*, *107*(2), 310–312. https://doi.org/10.1016/j.econlet.2010.02.020

Santos Silva, J. M. C., & Tenreyro, S. (2011). Further simulation evidence on the performance of the Poisson pseudo-maximum likelihood estimator. *Economics Letters*, *112*(2), 220–222. https://doi.org/10.1016/j.econlet.2011.05.008

Sarker, R., & Jayasinghe, S. (2007). Regional trade agreements and trade in agri-food products: Evidence for the European Union from gravity modeling using disaggregated data. *Agricultural Economics*, *37*(1), 93–104. https://doi.org/10.1111/j.1574-0862.2007.00227.x

Sartorius, K., & Kirsten, J. (2007). A framework to facilitate institutional arrangements for smallholder supply in developing countries: An agribusiness perspective. *Food Policy*, *32*(5), 640–655. https://doi.org/10.1016/j.foodpol.2007.03.001

Sauer, C. M., Reardon, T., Tschirley, D., Liverpool-Tasie, S., Awokuse, T., Alphonce, R., Ndyetabula, D., & Waized, B. (2021). Consumption of processed food & food away from home in big cities, small towns, and rural areas of Tanzania. *Agricultural Economics*, *52*(5), 749–770. https://doi.org/10.1111/agec.12652

Schipmann, C., & Qaim, M. (2011). Supply chain differentiation, contract agriculture, and farmers' marketing preferences: The case of sweet pepper in Thailand. *Food Policy*, *36*(5), 667–677. https://doi.org/10.1016/j.foodpol.2011.07.004

Sellare, J., Meemken, E.-M., Kouamé, C., & Qaim, M. (2020). Do Sustainability Standards Benefit Smallholder Farmers Also When Accounting For Cooperative Effects? Evidence from Côte d'Ivoire. *American Journal of Agricultural Economics*, *102*(2), 681–695. https://doi.org/10.1002/ajae.12015

Sen, A. (1981). Ingredients of Famine Analysis: Availability and Entitlements. *The Quarterly Journal of Economics*, *96*(3), 433–464. https://doi.org/10.2307/1882681

Shepherd, B., & Duval, Y. (2016). Estimating the benefits of cross-border paperless trade in Asia and the Pacific. In A. Saggu, C. Utoktham, & Y. Duval, *Reducing Trade Costs in AsiaPacific Developing Countries*. UNESCAP. https://www.unescap.org/sites/default/files/publications/STI84_Final.pdf

Shepherd, B., & Wilson, N. L. W. (2013). Product standards and developing country agricultural exports: The case of the European Union. *Food Policy*, *42*, 1–10. https://doi.org/10.1016/j.foodpol.2013.06.003

Silva, J. M. C. S., & Tenreyro, S. (2006). The Log of Gravity. *The Review of Economics and Statistics*, *88*(4), 641–658. https://doi.org/10.1162/rest.88.4.641

Simmons, P., Winters, P., & Patrick, I. (2005). An analysis of contract farming in East Java, Bali, and Lombok, Indonesia. *Agricultural Economics*, *33*(s3), 513–525. https://doi.org/10.1111/j.1574-0864.2005.00096.x

Simoes, A. J. G., & Hidalgo, C. A. (2012). *The Economic Complexity Observatory: An Analytical Tool for Understanding the Dynamics of Economic Development* [Thesis, Massachusetts Institute of Technology]. https://dspace.mit.edu/handle/1721.1/76532

Smith, L. C., & Frankenberger, T. R. (2018). Does Resilience Capacity Reduce the Negative Impact of Shocks on Household Food Security? Evidence from the 2014 Floods in Northern Bangladesh. *World Development*, *102*, 358–376. https://doi.org/10.1016/j.worlddev.2017.07.003

Soullier, G., & Moustier, P. (2018). Impacts of contract farming in domestic grain chains on farmer income and food insecurity. Contrasted evidence from Senegal. *Food Policy*, *79*, 179–198. https://doi.org/10.1016/j.foodpol.2018.07.004

Stein, A. J., & Qaim, M. (2007). The Human and Economic Cost of Hidden Hunger: *Food and Nutrition Bulletin*. https://doi.org/10.1177/156482650702800201

Stifel, D., & Minten, B. (2017). Market Access, Well-being, and Nutrition: Evidence from Ethiopia. *World Development*, *90*, 229–241. https://doi.org/10.1016/j.worlddev.2016.09.009

Stifel, D., Minten, B., & Koru, B. (2016). Economic Benefits of Rural Feeder Roads: Evidence from Ethiopia. *The Journal of Development Studies*, *52*(9), 1335–1356. https://doi.org/10.1080/00220388.2016.1175555

Stigler, G. J. (1972). Economic competition and political competition. *Public Choice*, *13*(1), 91–106. https://doi.org/10.1007/BF01718854

Stock, J. H., & Yogo, M. (2005). Testing for Weak Instruments in Linear IV Regression. In D. W. K. Andrews & J. H. Stock (Eds.), *Identification and Inference for Econometric Models: Essays in Honor of Thomas Rothenberg* (pp. 80–108). Cambridge University Press. https://doi.org/10.1017/CBO9780511614491.006

Sun, D., Liu, Y., Grant, J., Long, Y., Wang, X., & Xie, C. (2021). Impact of food safety regulations on agricultural trade: Evidence from China's import refusal data. *Food Policy*, *105*, 102185. https://doi.org/10.1016/j.foodpol.2021.102185

Sun, L., & Reed, M. R. (2010). Impacts of Free Trade Agreements on Agricultural Trade Creation and Trade Diversion. *American Journal of Agricultural Economics*, *92*(5), 1351–1363.

Swinnen, J. F. M., & Maertens, M. (2007). Globalization, privatization, and vertical coordination in food value chains in developing and transition countries. *Agricultural Economics*, *37*(s1), 89–102. https://doi.org/10.1111/j.1574-0862.2007.00237.x

Swinnen, J., & Vandeplas, A. (2012). Rich Consumers and Poor Producers: Quality and Rent Distribution in Global Value Chains. *Journal of Globalization and Development*, *2*(2). https://doi.org/10.1515/1948-1837.1036

Tambo, J. A., Uzayisenga, B., Mugambi, I., Bundi, M., & Silvestri, S. (2020). Plant clinics, farm performance and poverty alleviation: Panel data evidence from Rwanda. *World Development*, *129*, 104881. https://doi.org/10.1016/j.worlddev.2020.104881

Tessmann, J. (2018). Governance and upgrading in South–South value chains: Evidence from the cashew industries in India and Ivory Coast. *Global Networks*, *18*(2), 264–284. https://doi.org/10.1111/glob.12165

Tessmann, J. (2020). Global value chains and policy practice: The making of linkages in the Ivorian cashew industry. *Competition & Change*, *24*(1), 26–43. https://doi.org/10.1177/1024529419877491

Timmer, P. (2002). Agriculture and economic development. In B. L. Gardner & G. C. Rausser, *Handbook of Agricultural Economics* (1st ed., Vol. 2, pp. 1487–1546). Elsevier.

Tinbergen, J. (1962). *Shaping the World Economy; Suggestions for an International Economic Policy*. Twentieth Century Fund. https://repub.eur.nl/pub/16826

Ton, G., Vellema, W., Desiere, S., Weituschat, S., & D'Haese, M. (2018). Contract farming for improving smallholder incomes: What can we learn from effectiveness studies? *World Development*, *104*, 46–64. https://doi.org/10.1016/j.worlddev.2017.11.015

Tschirley, D. L., Snyder, J., Dolislager, M., Reardon, T., Haggblade, S., Goeb, J., Traub, L., Ejobi, F., & Meyer, F. (2015). Africa's unfolding diet transformation: Implications for agrifood system employment. *Journal of Agribusiness in Developing and Emerging Economies*, *5*(2), 102–136. https://doi.org/10.1108/JADEE-01-2015-0003

Tschirley, D., Reardon, T., Dolislager, M., & Snyder, J. (2015). The Rise of a Middle Class in East and Southern Africa: Implications for Food System Transformation. *Journal of International Development*, *27*(5), 628–646. https://doi.org/10.1002/jid.3107

UN. (2022). *World Population Prospects 2022*. United Nations Department of Economic and Social Affairs. https://www.un.org/development/desa/pd/sites/www.un.org.development.desa.pd/files/wpp2022_summary_of_results.pdf

United Nations. (2019). *Digital and Sustainable Trade Facilitation: Global Report 2019*. UNESCAP. https://untfsurvey.org/report

Van den Broeck, G., Mardulier, M., & Maertens, M. (2021). All that is gold does not glitter: Income and nutrition in Tanzania. *Food Policy*, *99*, 101975. https://doi.org/10.1016/j.foodpol.2020.101975

Väth, S. J., Gobien, S., & Kirk, M. (2019). Socio-economic well-being, contract farming and property rights: Evidence from Ghana. *Land Use Policy*, *81*, 878–888. https://doi.org/10.1016/j.landusepol.2017.04.023

Venables, A. J. (2004). Small, remote and poor? *World Trade Review*, *3*(3), 453–457.

Venables, A. J. (2008). New economic geography. In P. Macmillan, *The new palgrave dictionary of economics*. Palgrave Macmillan.

von Braun, J. (1994). Production, Employment, and Income Effects of Commercialization of Agriculture. In J. von Braun & E. T. Kennedy, *Agricultural Commercialization, Economic Development, and Nutrition* (pp. 37–64). The Johns Hopkins University Press. https://ebrary.ifpri.org/digital/collection/p15738coll2/id/129374

von Braun, J. (1995). Agricultural commercialization: Impacts on income and nutrition and implications for policy. *Food Policy*, *20*(3), 187–202. https://doi.org/10.1016/0306-9192(95)00013-5

von Braun, J., & Kennedy, E. T. (Eds.). (1994). *Agricultural commercialization, economic development, and nutrition*. The Johns Hopkins University Press. https://doi.org/nodoi

Wang, H. H., Wang, Y., & Delgado, M. S. (2014). The Transition to Modern Agriculture: Contract Farming in Developing Economies. *American Journal of Agricultural Economics*, *96*(5), 1257–1271. https://doi.org/10.1093/ajae/aau036

Were, M. (2015). Differential effects of trade on economic growth and investment: A cross-country empirical investigation. *Journal of African Trade*, *2*(1), 71–85. https://doi.org/10.1016/j.joat.2015.08.002

Westerlund, J., & Wilhelmsson, F. (2011). Estimating the gravity model without gravity using panel data. *Applied Economics*, *43*(6), 641–649. https://doi.org/10.1080/00036840802599784

Wesz Junior, V. J. (2022). Soybean production in Paraguay: Agribusiness, economic change and agrarian transformations. *Journal of Agrarian Change*, *22*(2), 317–340. https://doi.org/10.1111/joac.12436

Wiggins, S., & Keats, S. (2013). *Leaping and Learning: Linking smallholders to markets in Africa* (Agriculture for Impact; p. 114). Imperial College and Overseas Development Institute, ODI. https://cgspace.cgiar.org/bitstream/handle/10568/35246/L%26L%20Link%20SF%20to%20Mkt%20final%20May%202013.pdf

Williamson, O. E. (1979). Transaction-Cost Economics: The Governance of Contractual Relations. *The Journal of Law and Economics*, *22*(2), 233–261. https://doi.org/10.1086/466942

Williamson, O. E. (2002). The Theory of the Firm as Governance Structure: From Choice to Contract. *Journal of Economic Perspectives*, *16*(3), 171–195. https://doi.org/10.1257/089533002760278776

Wooldridge, J. M. (2007). Inverse probability weighted estimation for general missing data problems. *Journal of Econometrics*, *141*(2), 1281–1301. https://doi.org/10.1016/j.jeconom.2007.02.002

Wooldridge, J. M. (2010). *Econometric Analysis of Cross Section and Panel Data* (Second Edition). The MIT Press. https://mitpress.mit.edu/books/econometric-analysis-cross-section-and-panel-data-second-edition

World Bank. (2008). *World Development Report 2008: Agriculture for Development*. The World Bank. https://doi.org/10.1596/978-0-8213-6807-7

World Bank. (2020a). *Trading for Development in the Age of Global Value Chains* (world; World Development Report 2020). World Bank. https://elibrary.worldbank.org/doi/pdf/10.1596/978-1-4648-1457-0

World Bank. (2020b). *World Development Indicators*. The World Bank Group. https://databank.worldbank.org/source/world-development-indicators

Xiong, B., & Beghin, J. (2014). Disentangling Demand-Enhancing and Trade-Cost Effects of Maximum Residue Regulations. *Economic Inquiry*, *52*(3), 1190–1203. https://doi.org/10.1111/ecin.12082

Yang, J. (2009). *Small and Medium Enterprises (SME) Adjustments to Information Technology (IT) in Trade Facilitation: The South Korean Experien* (Working Paper No. 61; ARTNeT Working Paper Series, p. 40). Asia-Pacific Research and Training Network on Trade (ARTNeT). https://www.econstor.eu/bitstream/10419/64260/1/601134532.pdf

Yasui, T., & Engman, M. (2009). The Role of Automation in Trade Facilitation. In *Overcoming Border Bottlenecks: The Costs and Benefits of Trade Facilitation* (pp. 141–172). OECD. https://doi.org/10.1787/9789264056954-6-en

Yotov, Y., Piermartini, R., Monteiro, J.-A., & Larch, M. (2016). *An advanced guide to trade policy analysis: The structural gravity model*. World Trade Organization (WTO). https://www.wto.org/english/res_e/publications_e/advancedguide2016_e.htm

Zaki, C. (2010). *Does Trade Facilitation Matter in Bilateral Trade?* (Working Paper No. 515; Working Paper Series). Economic Research Forum. https://mpra.ub.uni-muenchen.de/23353/1/MPRA_paper_23353.pdf

Zaki, C. (2015). How Does Trade Facilitation Affect International Trade? *The European Journal of Development Research*, *27*(1), 156–185. https://doi.org/10.1057/ejdr.2014.36

7 Appendix

7.1 Appendix for Chapter 2

7.1.1 Tables

Table 2A.1: Types of shocks each household suffered

Type of shock	Households	Households %
Health	1,807	40.10
Natural	2,404	53.35
Economic	3,041	67.49
Crime/conflict	431	9.57
Number of shocks suffered	**Households**	**Households %**
None	754	16.73
One	1,108	24.59
Two	1,525	33.84
Three	951	21.11
All four	168	3.73

Note: Households do not add up to 100%; based on N = 4,506 households.

Table 2A.2: Transition probabilities (in %) across VML groups

VML	Cooperatives		VML	Processors	
Cooperatives	**N**	**Y**	**Processors**	**N**	**Y**
N	92.3	7.7	**N**	58.4	41.6
Y	51.4	48.6	**Y**	41.2	58.8
Total	89.5	10.5		54.3	45.7

Table 2A.3: Effect of vertical market linkages on farm-level outcomes

	Yield (log)				Commercialization			
	(1) OLS	(2) FD	(3) FE	(4) CRE	(5) OLS	(6) FD	(7) FE	(8) CRE
VML1	0.2416*	0.0195	0.0132	0.1130	0.1243***	0.0384	0.0423	0.0521*
	(0.1407)	(0.1191)	(0.1633)	(0.1575)	(0.0241)	(0.0234)	(0.0306)	(0.0288)
VML2	0.6614***	0.3769***	0.2832***	0.2997***	0.1206***	0.0593***	0.0571***	0.0595***
	(0.0858)	(0.0734)	(0.0808)	(0.0799)	(0.0146)	(0.0140)	(0.0160)	(0.0158)
VML1 x VML2	-0.8783***		-0.2956	-0.5453**	-0.0830**		-0.0472	-0.0720**
	(0.2342)		(0.2455)	(0.2262)	(0.0344)		(0.0392)	(0.0328)
$VML1_m$				-0.3353				0.1225***
				(0.2493)				(0.0432)
$VML2_m$				0.6359***				0.1529***
				(0.1806)				(0.0307)
Wave 2	0.4207***		0.8517***	0.7310***	0.0401*		0.0841***	0.0827***
	(0.1240)		(0.1164)	(0.1104)	(0.0219)		(0.0228)	(0.0214)
Wave 3	0.1741		0.6176***	0.4937***	-0.0109		0.0340	0.0324
	(0.1249)		(0.1174)	(0.1114)	(0.0220)		(0.0235)	(0.0221)
Off-farm labor	0.3154***	0.4351***	-0.0963	0.0557	-0.0492***	0.0049	-0.0593***	-0.0582***
	(0.1099)	(0.0836)	(0.1123)	(0.1008)	(0.0184)	(0.0168)	(0.0208)	(0.0179)
Farm size (acres)	0.0238***	-0.0528***	-0.0530***	-0.0530***	0.0106***	-0.0012	0.0005	0.0005
	(0.0084)	(0.0121)	(0.0115)	(0.0116)	(0.0014)	(0.0023)	(0.0021)	(0.0021)
Intercropping	1.9477***	0.7299***	0.9143***	0.9158***	0.0274*	0.0190	0.0526***	0.0526***
	(0.0912)	(0.0879)	(0.0896)	(0.0897)	(0.0149)	(0.0167)	(0.0177)	(0.0178)
Constant	0.2937***	0.0108	1.6470***	-0.8787***	0.3414***	-0.0188*	0.4017***	0.2553***
	(0.0991)	(0.0495)	(0.1467)	(0.1706)	(0.0241)	(0.0105)	(0.0292)	(0.0367)
Household FE	NO	NO	YES	YES	NO	NO	YES	YES
Time FE	NO	NO	YES	YES	NO	NO	YES	YES
Observations	4,506	3,004	4,506	4,506	4,506	3,004	4,506	4,506
R^2	0.252	0.100	0.128	0.316	0.120	0.057	0.075	0.193
AIC	19,912.9	13,412. 8	15,405.1		4,055.5	3,823.7	868.5	

Note: The table shows coefficient estimates with standard errors clustered at the household level in parentheses. Significance levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. For panel models, the overall R^2 is reported.

Table 2A.4: Effect of vertical market linkages on intermediate welfare

	Income (log)				Poverty			
	(1) OLS	(2) FD	(3) FE	(4) CRE	(5) OLS	(6) FD	(7) FE	(8) CRE
VML1	1.0930***	0.4835***	0.5293***	0.7178***	-0.0613**	-0.0323	-0.0437	-0.0488
	(0.1415)	(0.1606)	(0.2044)	(0.1993)	(0.0290)	(0.0301)	(0.0362)	(0.0345)
VML2	0.8871***	0.5403***	0.5303***	0.5672***	-0.0730***	-0.0396***	-0.0671***	-0.0665***
	(0.0852)	(0.0985)	(0.1122)	(0.1099)	(0.0148)	(0.0149)	(0.0167)	(0.0166)
VML1 x VML2	-1.0256***		-0.4265	-0.9004***	0.0303		0.0190	0.0311
	(0.1959)		(0.2781)	(0.2051)	(0.0477)		(0.0537)	(0.0476)
$VML1_m$				0.6499**				-0.0328
				(0.2814)				(0.0457)
$VML2_m$				0.8246***				-0.0195
				(0.2031)				(0.0291)
Wave 2	0.9524***		1.3841***	1.2326***	0.0572***		0.0227	0.0479**
	(0.1843)		(0.2024)	(0.1846)	(0.0194)		(0.0215)	(0.0193)
Wave 3	1.2626***		1.7098***	1.5526***	-0.0072		-0.0451**	-0.0196
	(0.1901)		(0.2074)	(0.1923)	(0.0203)		(0.0221)	(0.0199)
Off-farm labor	0.4669***	0.7804***	0.2069	0.3936***	-0.0889***	-0.0025	-0.0499**	-0.0827***
	(0.1456)	(0.1368)	(0.1720)	(0.1418)	(0.0172)	(0.0164)	(0.0208)	(0.0173)
Farm size (acres)	0.0741***	0.0793***	0.0709***	0.0709***	-0.0023	-0.0057*	-0.0054**	-0.0054**
	(0.0089)	(0.0229)	(0.0197)	(0.0200)	(0.0017)	(0.0031)	(0.0028)	(0.0028)
Intercropping	1.3298***	0.9228***	1.3211***	1.3231***	-0.0222	-0.0301*	-0.0418**	-0.0420**
	(0.1115)	(0.1279)	(0.1458)	(0.1461)	(0.0140)	(0.0167)	(0.0172)	(0.0172)
Constant	6.6493***	0.5359***	6.5216***	6.6265***	0.9319***	-0.0501***	0.9502***	0.9157***
	(0.3177)	(0.0805)	(0.3339)	(0.4391)	(0.0181)	(0.0109)	(0.0240)	(0.0325)
Household FE	NO	NO	YES	YES	NO	NO	YES	YES
Time FE	NO	NO	YES	YES	NO	NO	YES	YES
Observations	4,506	3,004	4,506	4,506	4,506	3,004	4,506	4,506
R^2	0.319	0.218	0.297	0.322	0.080	0.018	0.056	0.089
AIC	21,688.6	15,840.2	19,046.6		3,708.2	3,787.3	834.1	

Note: The table shows coefficient estimates with standard errors clustered at the household level in parentheses. Significance levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. For panel models, the overall R^2 is reported.

Table 2A.5: Effect of vertical market linkages on final welfare

	Food insecurity				Dietary diversity				Subjective well-being			
	(1) OLS	(2) FD	(3) FE	(4) CRE	(1) OLS	(2) FD	(3) FE	(4) CRE	(1) OLS	(2) FD	(3) FE	(4) CRE
VML1	-3.547***	-1.289	-2.418	-2.852**	0.290*	0.253	0.292	0.345*	-0.020	0.181*	0.142	0.095
	(1.063)	(1.334)	(1.571)	(1.435)	(0.155)	(0.169)	(0.210)	(0.193)	(0.100)	(0.104)	(0.130)	(0.124)
VML2	-2.255***	-2.169**	-2.438**	-2.714***	0.261***	0.247***	0.256***	0.272***	-0.035	0.101*	0.043	0.013
	(0.633)	(0.879)	(0.959)	(0.894)	(0.074)	(0.079)	(0.083)	(0.081)	(0.051)	(0.056)	(0.061)	(0.059)
VML1 x VML2	2.211		2.732	2.361	0.015		-0.094	-0.031	0.193		0.066	0.171
	(1.669)		(2.048)	(1.638)	(0.221)		(0.245)	(0.201)	(0.163)		(0.173)	(0.151)
$VML1_m$				-2.312				-0.070				-0.189
				(2.257)				(0.290)				(0.166)
$VML2_m$				-1.121				-0.005				-0.141
				(1.701)				(0.190)				(0.110)
Wave 2									0.455***		0.264***	0.360***
									(0.081)		(0.094)	(0.081)
Wave 3	0.673		0.424	0.403	-0.244***		-0.221***	-0.216***	0.521***		0.303***	0.404***
	(0.591)		(0.534)	(0.535)	(0.067)		(0.052)	(0.052)	(0.083)		(0.095)	(0.082)
Off-farm labor	-0.765	-1.329	-1.355	-1.052	-0.582***	0.068	0.068	-0.319***	-0.334***	-0.034	-0.175*	-0.271***
	(1.057)	(1.784)	(1.778)	(1.101)	(0.105)	(0.148)	(0.148)	(0.106)	(0.069)	(0.066)	(0.089)	(0.071)
Farm size (acres)	-0.469***	-0.518**	-0.516**	-0.621***	-0.032**	-0.021	-0.021	-0.020	-0.020**	-0.014	-0.020	-0.026*
	(0.110)	(0.218)	(0.217)	(0.188)	(0.014)	(0.022)	(0.022)	(0.018)	(0.008)	(0.015)	(0.015)	(0.014)
Intercropping	1.890***	-0.319	-0.302	0.134	-0.242***	-0.040	-0.040	-0.043	-0.036	-0.053	-0.053	-0.044
	(0.646)	(1.051)	(1.047)	(1.001)	(0.072)	(0.096)	(0.096)	(0.093)	(0.051)	(0.060)	(0.062)	(0.060)
Constant	8.817***	0.396	10.313***	8.964***	7.623***	-0.220***	7.097***	7.866***	3.988***	0.061	3.987***	4.036***
	(1.484)	(0.535)	(2.412)	(1.891)	(0.152)	(0.052)	(0.195)	(0.235)	(0.094)	(0.040)	(0.109)	(0.135)
Household FE	NO	NO	YES	YES	NO	NO	YES	YES	NO	NO	YES	YES
Time FE	NO	NO	YES	YES	NO	NO	YES	YES	NO	NO	YES	YES
Observations	3,004	1,502	3,004	3,004	3,004	1,502	3,004	3,004	4,013	2,432	4,013	4,013
R^2	0.053	0.027	0.028	0.091	0.070	0.019	0.029	0.119	0.038	0.015	0.025	0.066
AIC	25,144.5	13,328.5	22,456.6		12,152.3	6,342.4	8,486.1		13,491.0	9,105.1	10,610.4	

Note: The table shows coefficient estimates with standard errors clustered at the household level in parentheses. Significance levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. For panel models, the overall R^2 is reported.

Table 2A.6: Summarized balancing statistics after matching

	VML1			VML2		
Matching technique / Balancing statistics	**Nearest-neighbor**	**Radius**	**Kernel**	**Nearest-neighbor**	**Radius**	**Kernel**
Pseudo R2	0.017	0.004	0.012	0.010	0.003	0.002
LR chi2	11.32	2.71	8.88	31.56	8.36	7.25
P>chi2	1.000	1.000	1.000	0.539	1.000	1.000
Mean of standardized bias (%)	3.5	1.7	3.8	3.2	1.4	1.5
Median of standardized bias (%)	2.6	1.2	3.1	2.3	1.0	1.3
Rubin's B	29.6*	14.9	23.9	23.2	11.9	11.1
Rubin's R	0.97	0.71	0.55	0.98	0.90	1.09

Note: * indicates if Rubin's B > 25% threshold. ** indicates if Rubin's R is outside the [0.5;2] threshold.

Table 2A.7: Average treatment effect on the treated of the outcome variables, VML1

Outcome variable / Matching technique	**Nearest-neighbor**		**Radius**		**Kernel**	
	ATT	Critical hidden bias	ATT	Critical hidden bias	ATT	Critical hidden bias
Income (log)	0.537*** (0.148)	1.4 – 1.5	0.379*** (0.099)	1.6 – 1.7	0.395*** (0.107)	1.7 – 1.8
Yield (log)	0.282** (0.122)	> 2	0.232** (0.091)	1.2 – 1.3	0.182** (0.093)	1.1 – 1.2
Poverty	-0.033 (0.039)	< 1	-0.003 (0.029)	1.3 – 1.4	-0.022 (0.032)	< 1
Commercialization	0.081** (0.033)	1.1 – 1.2	0.091*** (0.022)	1.6 – 1.7	0.092*** (0.024)	1.8 – 1.9
Subjective well-being	0.085 (0.135)	< 1	0.060 (0.103)	< 1	0.067 (0.106)	< 1
HFIAS	-0.394 (1.404)	< 1	-1.988 (1.128)	< 1	-3.408** (1.373)	> 2
HDDS	0.099 (0.200)	1.0 – 1.1	0.024 (0.140)	< 1	0.036 (0.166)	< 1

Note: The table shows coefficient estimates with standard errors clustered at the household level in parentheses. Significance levels: * p < 0.10, ** p < 0.05, *** p < 0.01.

Table 2A.8: Average treatment effect on the treated of the outcome variables, VML2

Outcome variable / Matching technique	**Nearest-neighbor**		**Radius**		**Kernel**	
	ATT	Critical hidden bias	ATT	Critical hidden bias	ATT	Critical hidden bias
Income (log)	0.754*** (0.096)	> 2	0.667*** (0.065)	> 2	0.671*** (0.064)	> 2
Yield (log)	0.299*** (0.068)	1.4 – 1.5	0.277*** (0.054)	1.7 – 1.8	0.283*** (0.053)	1.8 – 1.9
Poverty	-0.066*** (0.021)	1.2 – 1.3	-0.080*** (0.017)	< 1	-0.077*** (0.016)	< 1
Commercialization	0.073*** (0.022)	1.2 – 1.3	0.088*** (0.016)	1.5 – 1.6	0.089*** (0.015)	1.5 – 1.6
Subjective well-being	0.020 (0.075)	< 1	0.015 (0.060)	< 1	0.019 (0.058)	< 1
HFIAS	-0.878 (0.898)	1.1 – 1.2	-1.536** (0.705)	> 2	-1.619** (0.691)	> 2
HDDS	0.199* (0.108)	1.0 – 1.1	0.267*** (0.082)	1.4 – 1.5	0.275** (0.081)	1.3 – 1.4

Note: The table shows coefficient estimates with standard errors clustered at the household level in parentheses. Significance levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

7.1.2 Figures

Source of Figures 2A.1 – 2A.9: Author's own elaboration based on NPS-LSMS data.

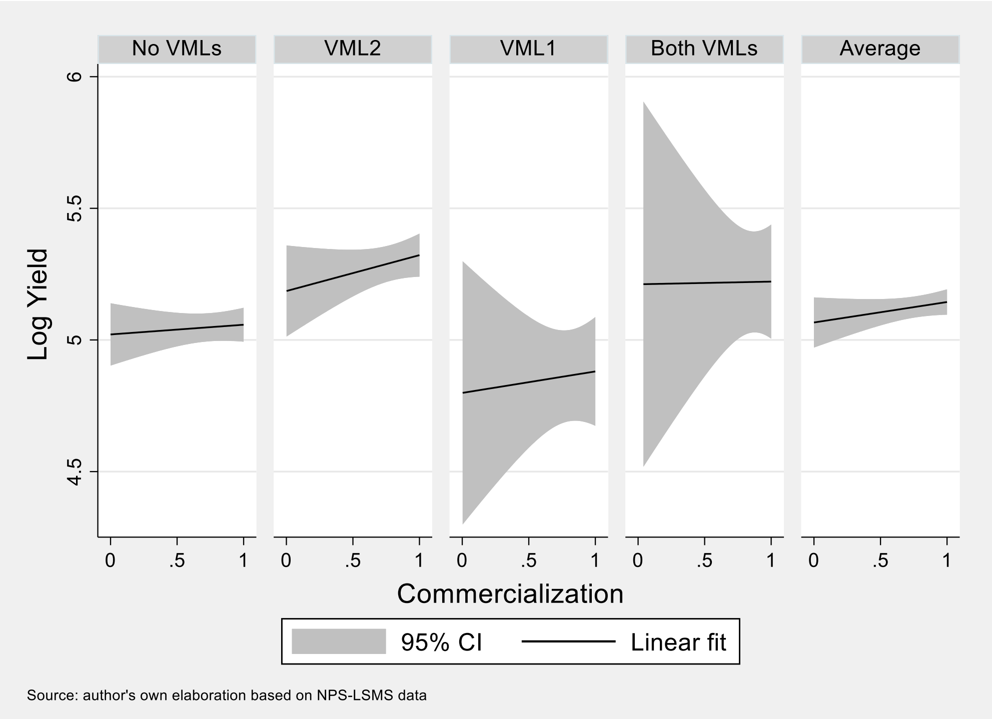


Figure 2A.1: Log yield vs. commercialization

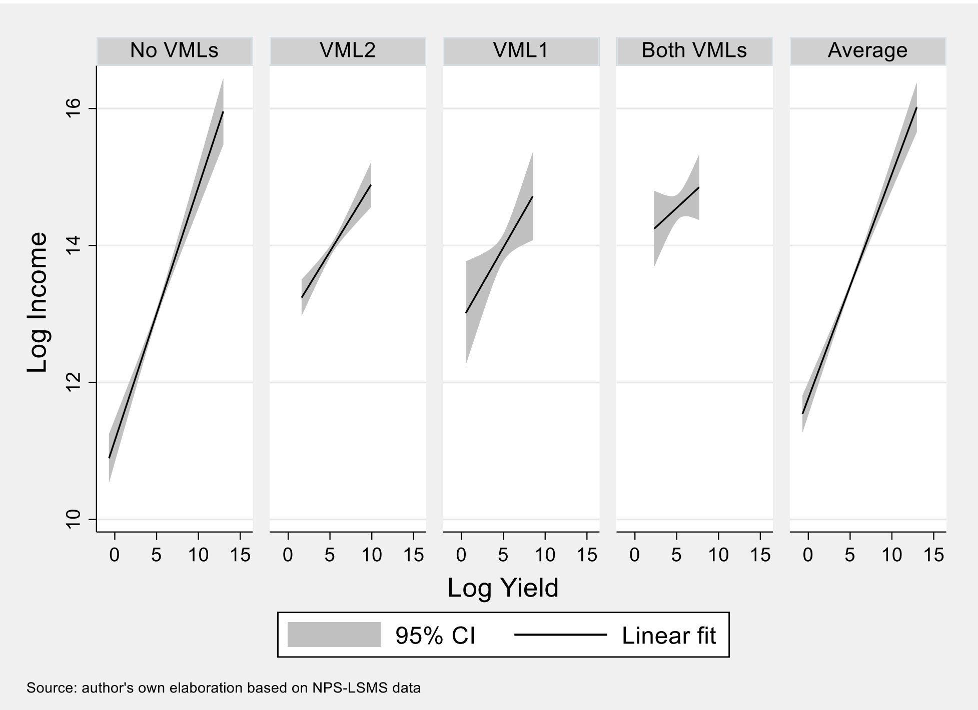


Figure 2A.2: Log income vs. log yield

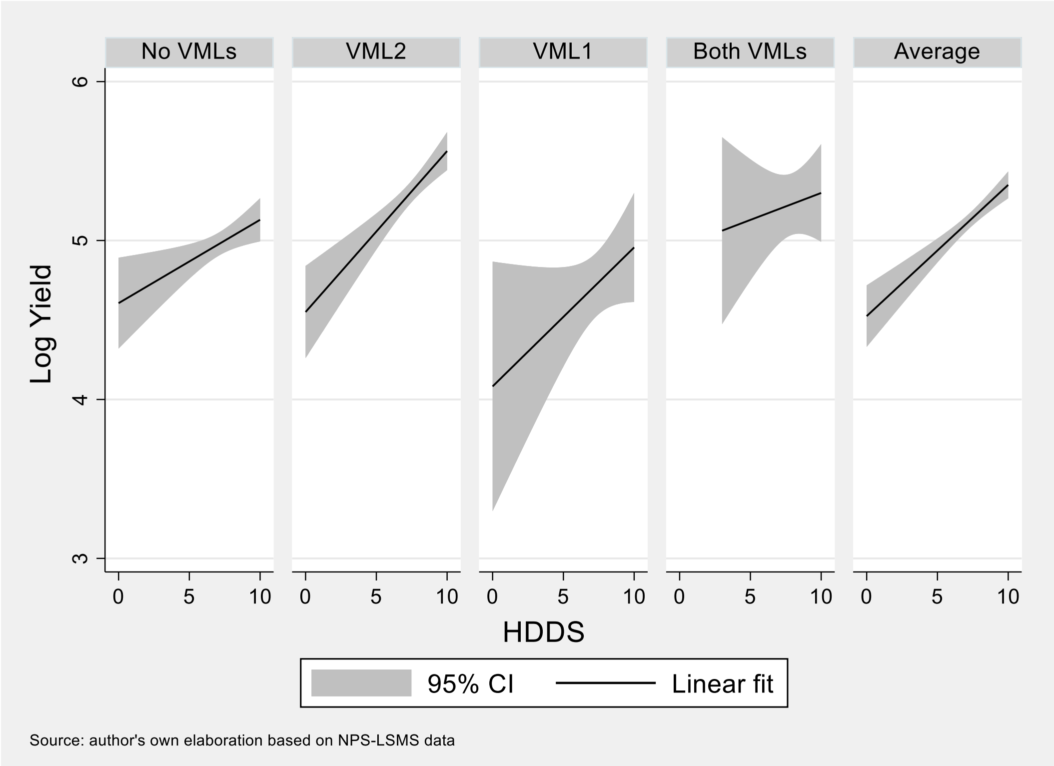


Figure 2A.3: Log yield vs. dietary diversity

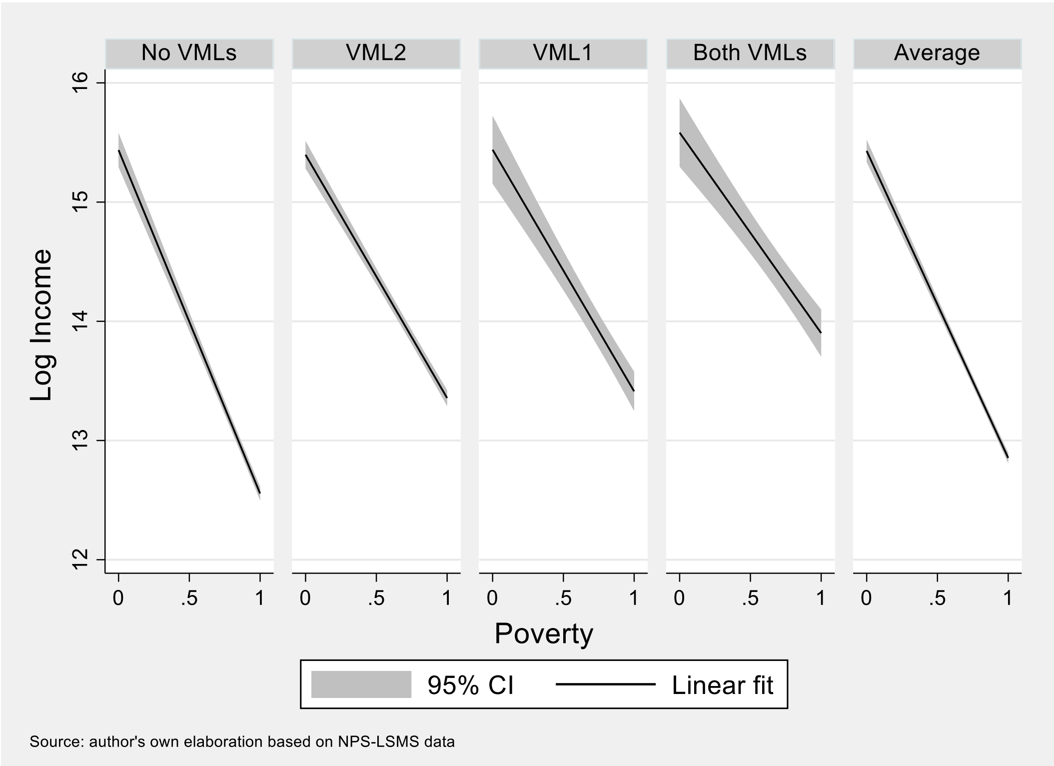


Figure 2A.4: Log income vs. poverty

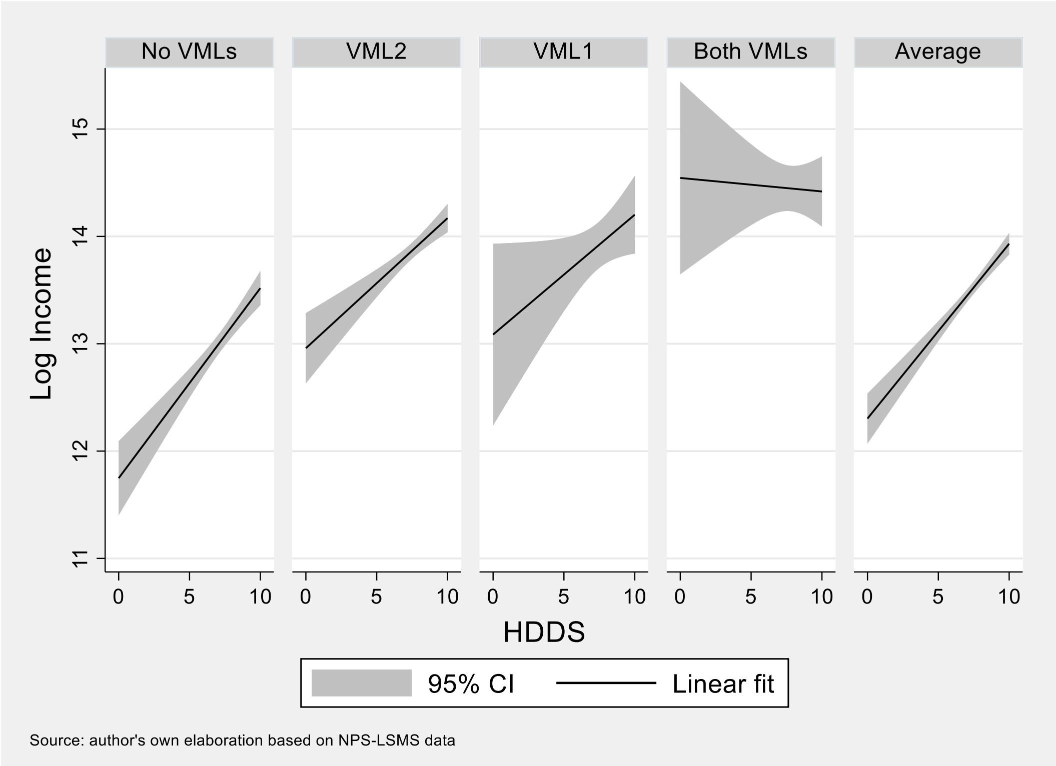


Figure 2A.5: Log income vs. dietary diversity

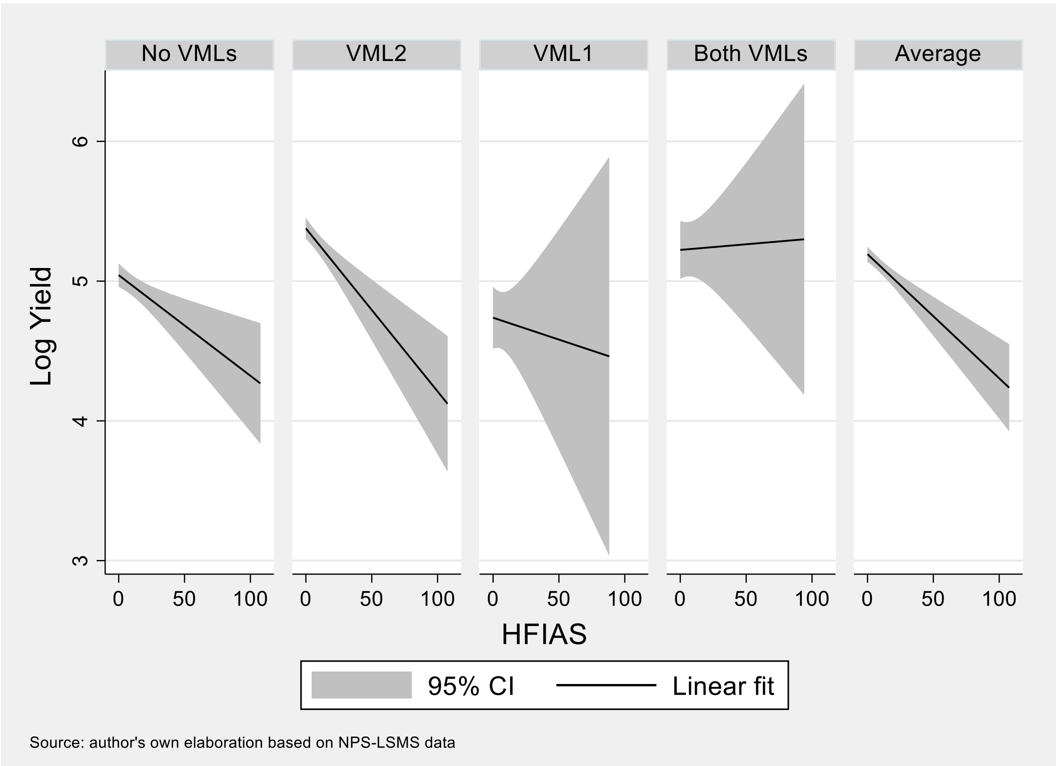


Figure 2A.6: Log yield vs. food insecurity

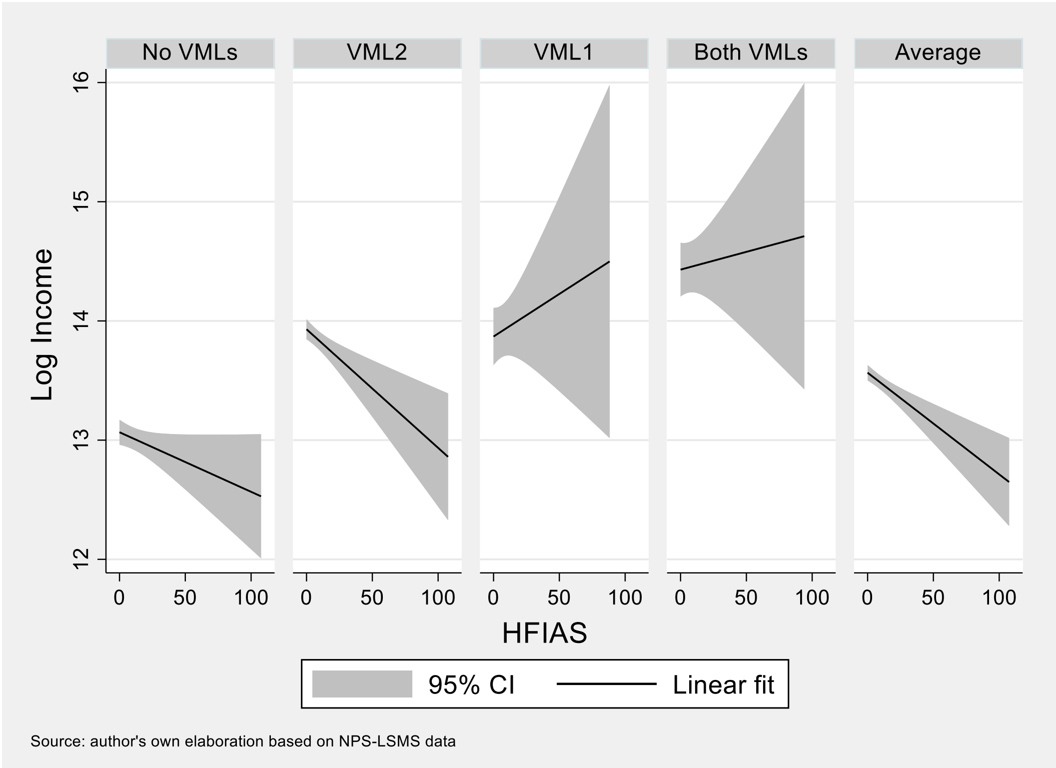


Figure 2A.7: Log income vs. food insecurity

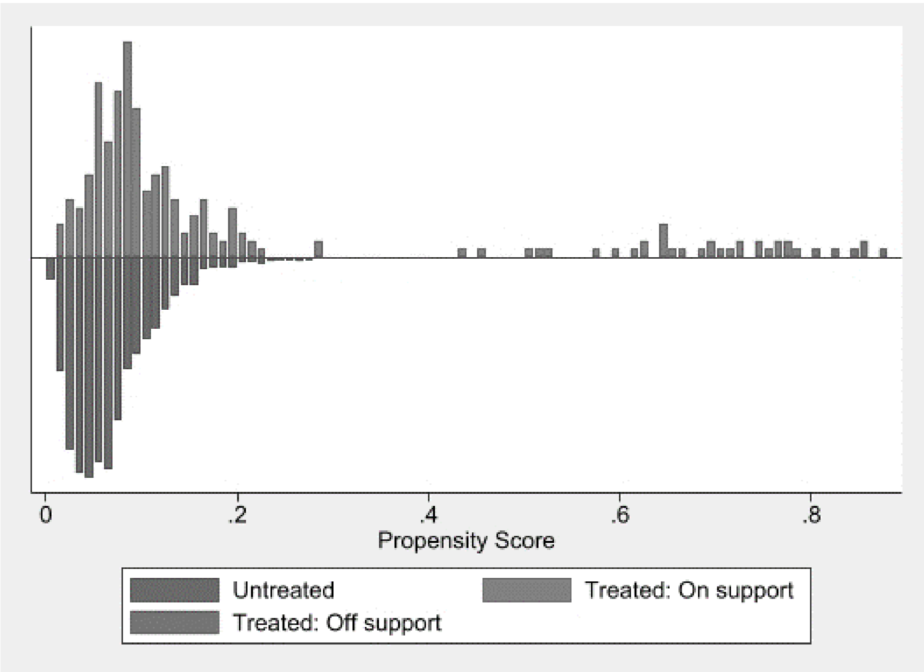


Figure 2A.8: Common support for VML1: Propensity scores using kernel matching

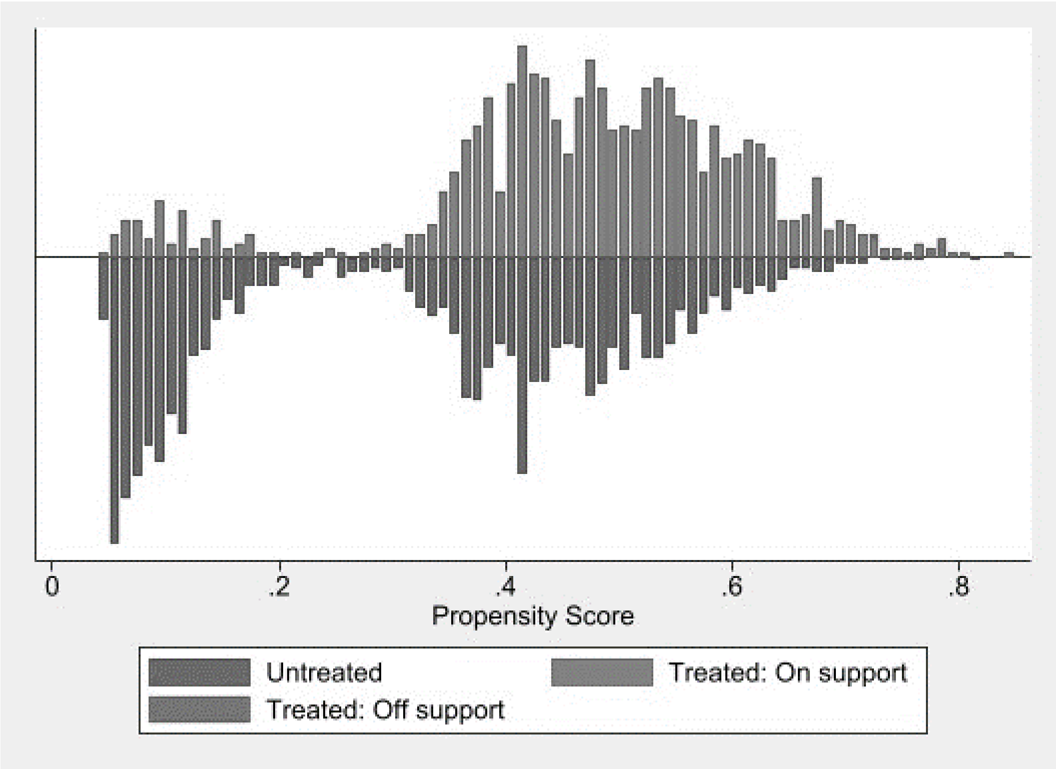


Figure 2A.9: Common support for VML2: Propensity scores using kernel matching

7.2 Appendix for Chapter 3

7.2.1 Sample size estimation

In order to ensure statistical power, we use the following formula to estimate our sample size (n) (Wooldridge, 2010):

$$n = \frac{z^2 \times p(1-p)}{\varepsilon^2} \quad \text{(A1),}$$

from which z is the standardized z-score considering a confidence level of 95%, p is the population proportion, and ε is the associated margin of error (5%). Since the true proportion of contracting farmers in the total population of cashew nuts farmers in Tanzania is unknown due to unavailability of precise data, we chose p = 0.5 *ex-ante* and ended up the data collection with p = 0.34 *ex-post*. This means that our target sample size was 384 and we reduced this target to 323 as we adopted a non-proportional random sampling strategy (see more details in Section 3.3.2), with a final sample size of 339 farmers' households.

7.2.2 Figures

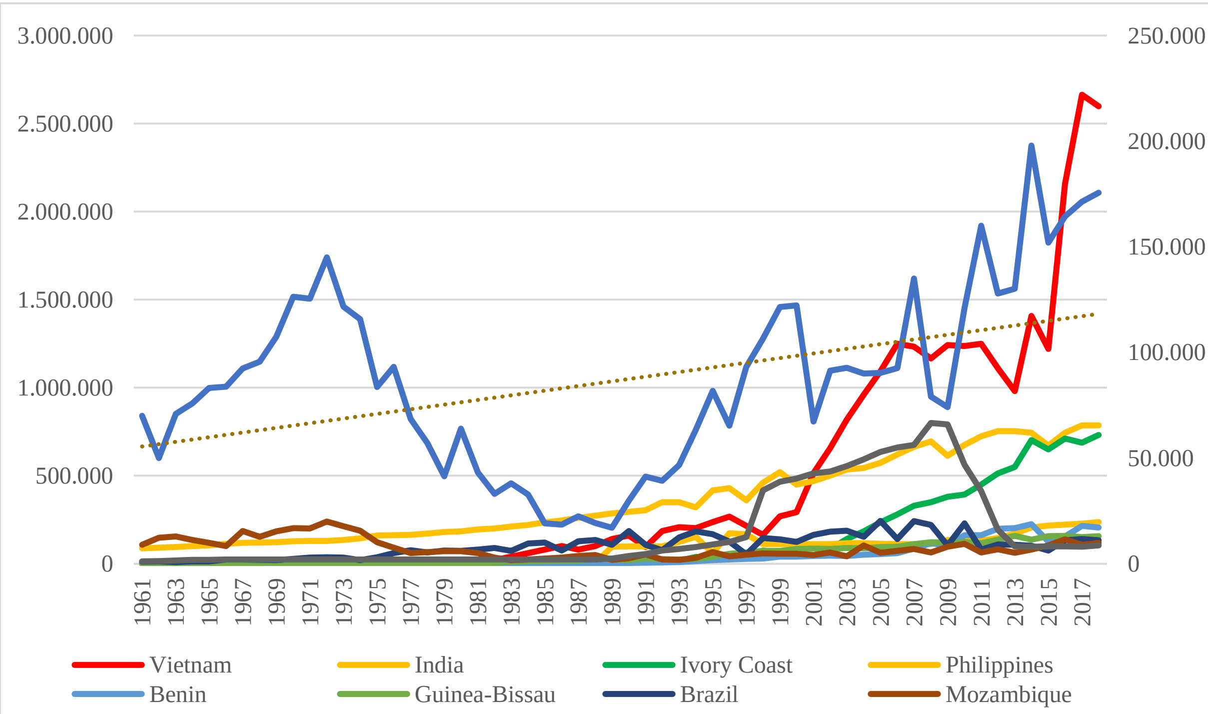


Figure 3A.1: Long trend of cashew nuts production in Tanzania

Note: Left axis: MT world output. Right axis: MT Tanzania output. Source: Author's own elaboration based on FAOSTAT data.

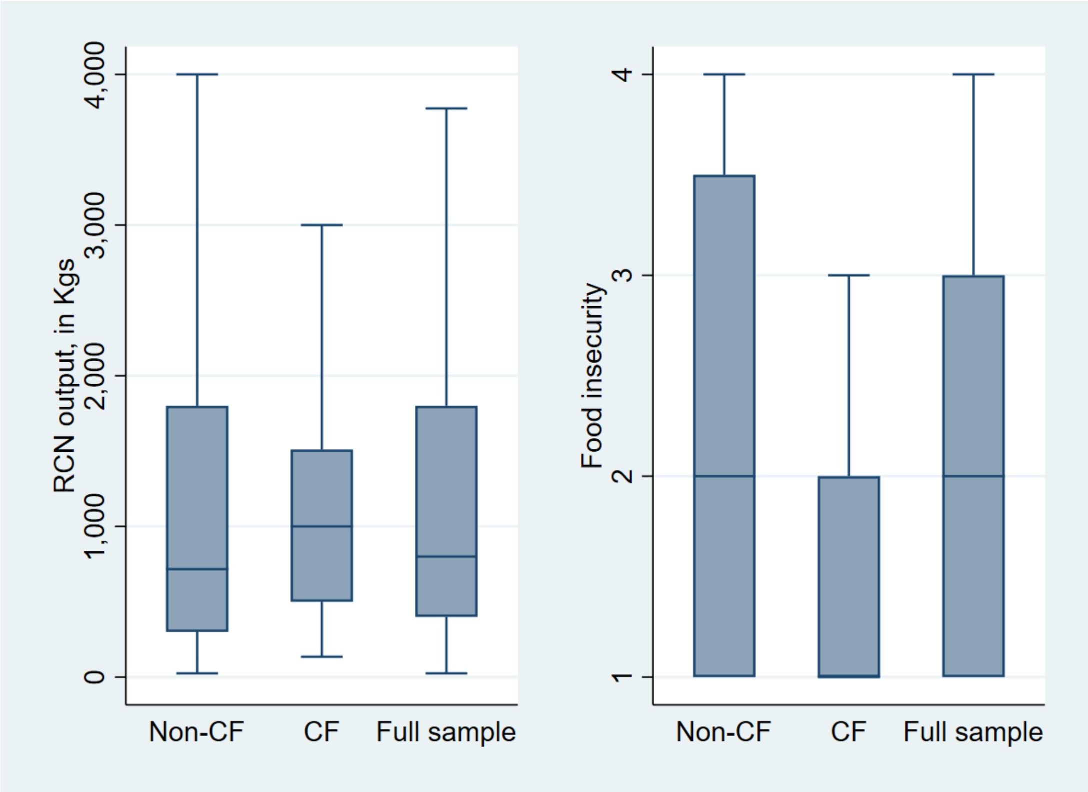


Figure 3A.2: Average of outcome variables by CF status

Source: Author's own elaboration based on fieldwork primary data.

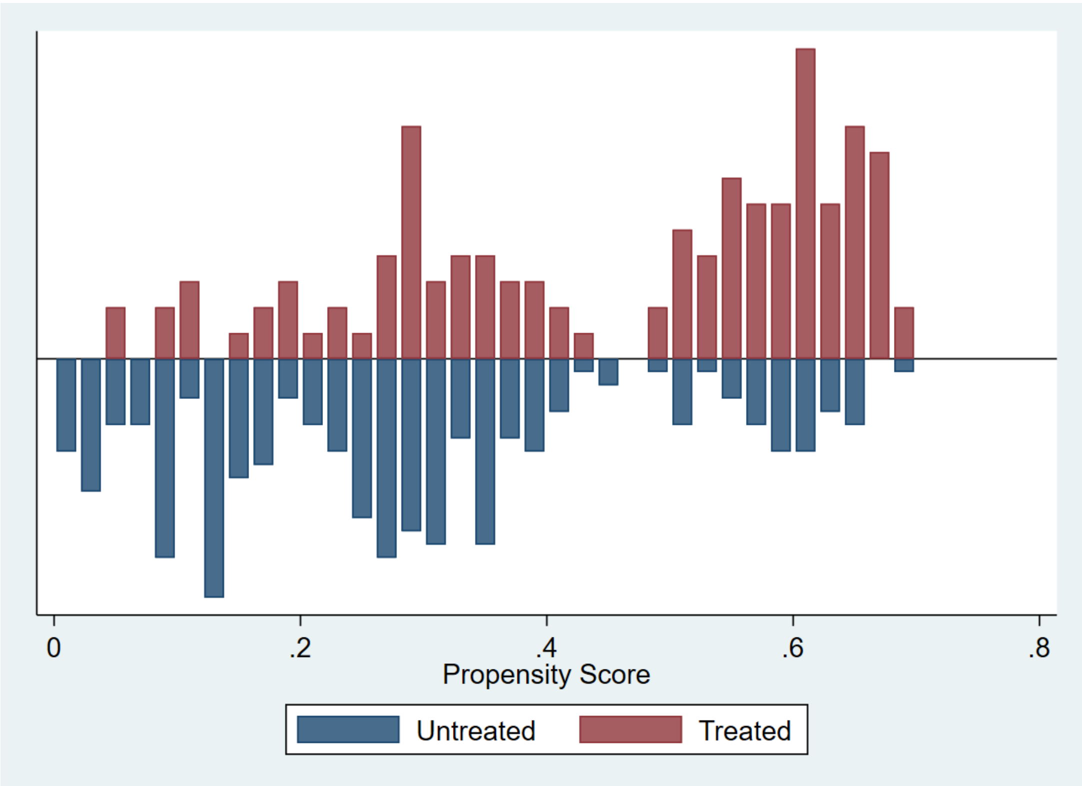


Figure 3A.3: Common support graph for matching variables

Source: Author's own elaboration based on fieldwork primary data.

7.2.3 Tables

Table 3A.1: Correlation matrix: instruments and contract farming (CF)

	CF (1/0)
CF (1/0)	1.000
Distance (in Kms) to the nearest CF farmer	-0.817***
Number of CF farmers in a 10 Km radius	0.959***
Number of CF farmers in a 25 Km radius	0.959***
Number of CF farmers in a 100 Km radius	0.348***
Number of CF farmers in a 200 Km radius	-0.0082

Note: *** significant at the 1% level with Bonferroni-corrected p-value.

Table 3A.2: Falsification test: instruments on outcome variables

	(1) First-stage IV CF participation (1/0)	(2) RCN output	(3) Food insecurity (HFIAS)
Distance to the nearest CF farmer (in Km)	-0.01083***	-33.43	0.07
	(0.00)	(52.79)	(0.06)
Distance to the nearest CF farmer (in Km2)	0.00003***	0.03	-0.00
	(0.00)	(0.13)	(0.00)
CF farmers in 10km radius	0.00312***	-37.36	0.08
	(0.00)	(44.82)	(0.05)
CF farmers in 100km radius	-0.00334***	-22.89	0.02
	(0.00)	(17.21)	(0.02)
HH size	0.00	61.19	-0.01
	(0.00)	(41.93)	(0.02)
Age of HH head	-0.00*	-15.31*	0.00
	(0.00)	(7.55)	(0.01)
Years of education of HH head	0.00	12.66	-0.05
	(0.00)	(23.32)	(0.03)
Gender of HH head (Female = 1)	-0.01	-69.05	0.08
	(0.01)	(147.18)	(0.17)
Access to credit (0/1)	0.01*	-11.59	0.00
	(0.01)	(283.33)	(0.21)
HH is not migrant (0/1)	-0.02	207.40	-0.20*
	(0.02)	(155.53)	(0.11)
Farm land under cashew nuts (acres)	0.00	13.80	-0.02
	(0.00)	(8.57)	(0.01)
Family labor (days per year)	0.00	-0.37	0.00***
	(0.00)	(0.58)	(0.00)
Hired labor (days per year)	0.00	3.60	-0.01**
	(0.00)	(2.51)	(0.00)
Application of herbicide/pesticide/fungicide	0.00	1.78**	-0.00

(litres)			
	(0.00)	(0.80)	(0.00)
Soil quality good or excellent (0/1)	0.01	128.13	-0.38***
	(0.01)	(128.54)	(0.11)
Soil steep (0/1)	0.01	-139.92	-0.14
	(0.01)	(353.93)	(0.28)
Received agricultural training (0/1)	0.00	-93.69	-0.27*
	(0.01)	(118.95)	(0.14)
District term	0.00	236.90	-0.34
	(0.02)	(193.86)	(0.24)
Constant	1.01***	4614.25	-2.22
	(0.19)	(4842.08)	(5.71)
Observations	339	222	222
First-stage F statistic	4,177.21***		
First-stage R2	0.975		

Note: Standard errors in parentheses: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 3A.3: Falsification test: instruments on outcome variables

	(1) First-stage IV CF participation (1/0)	(2) RCN output	(3) Food insecurity (HFIAS)
Distance to the nearest CF farmer (in Km)	-0.00939***	-33.43	0.07
	(0.00)	(52.79)	(0.06)
Distance to the nearest CF farmer (in Km2)	0.00003***	0.03	-0.00
	(0.00)	(0.13)	(0.00)
CF farmers in 10km radius	0.00437***	-37.36	0.08
	(0.00)	(44.82)	(0.05)
CF farmers in 200km radius	-0.00507***	-38.77	0.03
	(0.00)	(29.16)	(0.04)
HH size	0.00	61.19	-0.01
	(0.00)	(41.93)	(0.02)
Age of HH head	-0.00*	-15.31*	0.00
	(0.00)	(7.55)	(0.01)
Years of education of HH head	0.00	12.66	-0.05
	(0.00)	(23.32)	(0.03)
Gender of HH head (Female = 1)	-0.01	-69.05	0.08
	(0.01)	(147.18)	(0.17)
Access to credit (0/1)	0.01*	-11.59	0.00
	(0.01)	(283.33)	(0.21)
HH is not migrant (0/1)	-0.02	207.40	-0.20*
	(0.02)	(155.53)	(0.11)
Farm land under cashew nuts (acres)	0.00	13.80	-0.02
	(0.00)	(8.57)	(0.01)
Family labor (days per year)	0.00	-0.37	0.00***
	(0.00)	(0.58)	(0.00)
Hired labor (days per year)	0.00	3.60	-0.01**
	(0.00)	(2.51)	(0.00)
Application of herbicide/pesticide/fungicide (litres)	0.00	1.78**	-0.00
	(0.00)	(0.80)	(0.00)

Soil quality good or excellent (0/1)	0.01	128.13	-0.38***
	(0.01)	(128.54)	(0.11)
Soil steep (0/1)	0.01	-139.92	-0.14
	(0.01)	(353.93)	(0.28)
Received agricultural training (0/1)	0.00	-93.69	-0.27*
	(0.01)	(118.95)	(0.14)
District term	-0.00	236.90	-0.34
	(0.01)	(193.86)	(0.24)
Constant	1.01***	4614.25	-2.22
	(0.19)	(4842.08)	(5.71)
Observations	339	222	222
First-stage F statistic	4,744.15***		
First-stage R2	0.975		

Note: Standard errors in parentheses: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. OLS model.

Table 3A.4: Falsification test: instruments on outcome variables

	(1) First-stage IV CF participation (1/0)	(2) RCN output	(3) Food insecurity (HFIAS)
Distance to the nearest CF farmer (in Km)	-0.01083***	-33.43	0.07
	(0.00)	(52.79)	(0.06)
Distance to the nearest CF farmer (in Km2)	0.00003***	0.03	-0.00
	(0.00)	(0.13)	(0.00)
CF farmers in 25km radius	0.00312***	-37.36	0.08
	(0.00)	(44.82)	(0.05)
CF farmers in 100km radius	-0.00339***	-22.89	0.02
	(0.00)	(17.21)	(0.02)
HH size	0.00	61.19	-0.01
	(0.00)	(41.93)	(0.02)
Age of HH head	-0.00*	-15.31*	0.00
	(0.00)	(7.55)	(0.01)
Years of education of HH head	0.00	12.66	-0.05
	(0.00)	(23.32)	(0.03)
Gender of HH head (Female = 1)	-0.01	-69.05	0.08
	(0.01)	(147.18)	(0.17)
Access to credit (0/1)	0.01*	-11.59	0.00
	(0.01)	(283.33)	(0.21)
HH is not migrant (0/1)	-0.02	207.40	-0.20*
	(0.02)	(155.53)	(0.11)
Farm land under cashew nuts (acres)	0.00	13.80	-0.02
	(0.00)	(8.57)	(0.01)
Family labor (days per year)	0.00	-0.37	0.00***
	(0.00)	(0.58)	(0.00)
Hired labor (days per year)	0.00	3.60	-0.01**
	(0.00)	(2.51)	(0.00)
Application of herbicide/pesticide/fungicide (litres)	0.00	1.78**	-0.00
	(0.00)	(0.80)	(0.00)
Soil quality good or excellent (0/1)	0.01	128.13	-0.38***

	(0.01)	(128.54)	(0.11)
Soil steep (0/1)	0.01	-139.92	-0.14
	(0.01)	(353.93)	(0.28)
Received agricultural training (0/1)	0.00	-93.69	-0.27*
	(0.01)	(118.95)	(0.14)
District term	0.00	236.90	-0.34
	(0.02)	(193.86)	(0.24)
Constant	1.06***	5932.57	-3.34
	(0.18)	(5808.75)	(6.88)
Observations	339	222	222
First-stage F statistic	4,177.21***		
First-stage R2	0.975		

Note: Standard errors in parentheses: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. OLS model.

Table 3A.5: Falsification test: instruments on outcome variables

	(1) First-stage IV CF participation (1/0)	(2) RCN output	(3) Food insecurity (HFIAS)
Distance to the nearest CF farmer (in Km)	-0.00939***	-33.43	0.07
	(0.00)	(52.79)	(0.06)
Distance to the nearest CF farmer (in Km2)	0.00003***	0.03	-0.00
	(0.00)	(0.13)	(0.00)
CF farmers in 25km radius	0.00437***	-37.36	0.08
	(0.00)	(44.82)	(0.05)
CF farmers in 200km radius	-0.00507***	-38.77	0.03
	(0.00)	(29.16)	(0.04)
HH size	0.00	61.19	-0.01
	(0.00)	(41.93)	(0.02)
Age of HH head	-0.00*	-15.31*	0.00
	(0.00)	(7.55)	(0.01)
Years of education of HH head	0.00	12.66	-0.05
	(0.00)	(23.32)	(0.03)
Gender of HH head (Female = 1)	-0.01	-69.05	0.08
	(0.01)	(147.18)	(0.17)
Access to credit (0/1)	0.01*	-11.59	0.00
	(0.01)	(283.33)	(0.21)
HH is not migrant (0/1)	-0.02	207.40	-0.20*
	(0.02)	(155.53)	(0.11)
Farm land under cashew nuts (acres)	0.00	13.80	-0.02
	(0.00)	(8.57)	(0.01)
Family labor (days per year)	0.00	-0.37	0.00***
	(0.00)	(0.58)	(0.00)
Hired labor (days per year)	0.00	3.60	-0.01**
	(0.00)	(2.51)	(0.00)
Application of herbicide/pesticide/fungicide (litres)	0.00	1.78**	-0.00
	(0.00)	(0.80)	(0.00)
Soil quality good or excellent (0/1)	0.01	128.13	-0.38***
	(0.01)	(128.54)	(0.11)
Soil steep (0/1)	0.01	-139.92	-0.14

	(0.01)	(353.93)	(0.28)
Received agricultural training (0/1)	0.00	-93.69	-0.27*
	(0.01)	(118.95)	(0.14)
District term	-0.00	236.90	-0.34
	(0.01)	(193.86)	(0.24)
Constant	1.06***	5932.57	-3.34
	(0.18)	(5808.75)	(6.88)
Observations	339	222	222
First-stage F statistic	4,744.15***		
First-stage R2	0.975		

Note: Standard errors in parentheses: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. OLS model.

Table 3A.6: Test statistics for the model using the IV-2SLS estimator

	Test of endogeneity		Overidentifying restrictions		Overidentification test	Underidentification test	Weak identification test
	F-stat	*P*-value	Chi^2	*P*-value	Hansen *J*-stat	Kleibergen-Paap *LM*-stat	Cragg-Donald Wald *F*-stat
(1) RCN output	0.373	0.542	4.037	0.258	4.875	160.09***	2,268.87***
(2) Food insecurity (HFIAS)	1.620	0.204	1.212	0.750	1.476	160.09***	2,268.87***
(3) RCN output	0.342	0.559	4.157	0.245	4.960	160.08***	2,283.80***
(4) Food insecurity (HFIAS)	1.650	0.200	1.221	0.748	1.474	160.08***	2,283.80***
(5) RCN output	0.373	0.531	4.037	0.258	4.875	160.09***	2,268.87***
(6) Food insecurity (HFIAS)	1.620	0.204	1.212	0.750	1.476	160.09***	2,268.87***
(7) RCN output	0.342	0.548	4.157	0.245	4.960	160.08***	2,283.80***
(8) Food insecurity (HFIAS)	1.650	0.200	1.221	0.748	1.474	160.08***	2,283.80***

Standard errors in parentheses: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Note: Models for each pair of outcome variables always include the IVs distance and squared distance to nearest CF farmer and are shown in sequence for the IV 'number of CF farmers within a X Km radius' considering the following radius: 10-100; 10-200; 25-100; 25-200.

Table 3A.7: Determinants of contract farming participation

	(1) Probit	(2) Marginal effect
HH size	0.057	0.019
	(0.047)	(0.016)
Age of HH head	-0.017***	-0.006**
	(0.005)	(0.003)
Years of education of HH head	-0.171***	-0.058***
	(0.029)	(0.015)
Gender of HH head (Female = 1)	-0.811***	-0.276***

	(0.192)	(0.087)
Access to credit (0/1)	-0.413**	-0.141*
	(0.194)	(0.076)
HH is not migrant (0/1)	-0.595**	-0.203
	(0.300)	(0.128)
Farm land under cashew nuts (acres)	-0.034*	-0.012*
	(0.020)	(0.006)
Family labor (days per year)	-0.001	-0.0004
	(0.001)	(0.0004)
Hired labor (days per year)	0.001	0.0004
	(0.002)	(0.0005)
Application of herbicide/pesticide/fungicide (litres)	0.000	0.0001
	(0.000)	(0.0001)
Soil quality good or excellent (0/1)	0.078	0.027
	(0.103)	(0.037)
Soil steep (0/1)	-0.018	-0.006
	(0.167)	(0.057)
Received agricultural training (0/1)	0.830***	0.283***
	(0.122)	(0.080)
District term	0.132	0.045
	(0.231)	(0.076)
Constant	1.545	
	(1.497)	
Observations	339	
Pseudo R-squared	0.243	

Note: Standard errors in parentheses: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 3A.8: Effects of contract farming on RCN output, full table with controls

	(1) OLS	(2) PSM-ATT	(3) IPWRA-ATE	(4) IPWRA-ATT	(5) IV
Contract farming	241.216**	318.451*	178.206*	358.738**	172.694**
	(112.455)	(171.747)	(104.878)	(116.331)	(76.621)
HH size	57.122*				53.716**
	(29.305)				(27.099)
Age of HH head	-10.861**				-11.872**
	(4.648)				(4.844)
Years of education of HH head	-12.454				-16.427
	(25.152)				(25.011)
Gender of HH head (Female = 1)	4.488				-27.024
	(126.441)				(119.479)
Access to credit (0/1)	-47.693				-43.820
	(214.999)				(233.206)
HH is not migrant (0/1)	-48.776				14.314
	(139.595)				(133.849)
Farm land under	20.929				14.948

cashew nuts (acres)				
	(13.463)			(9.763)
Family labor (days per year)	0.329			0.526
	(0.565)			(0.495)
Hired labor (days per year)	4.100*			3.671
	(2.267)			(2.262)
Application of herbicide/pesticide /fungicide (litres)	2.414***			2.799***
	(0.666)			(0.543)
Soil quality good or excellent (0/1)	144.645			156.705*
	(89.274)			(84.091)
Soil steep (0/1)	-137.143			-178.164
	(310.736)			(329.544)
Received agricultural training (0/1)	-79.099			-53.552
	(109.998)			(103.521)
District term	-11.739			-18.221
	(41.501)			(33.061)
Constant	835.263**			887.970**
	(397.646)			(357.048)
Potential outcome mean				
0. No		1167.417***	1163.134***	
		(126.782)	(104.730)	
OME0				
HH size		41.475	13.955	
		(28.900)	(26.590)	
Age of HH head		-12.902**	-6.710	
		(5.195)	(5.314)	
Years of education of HH head		-4.440	-44.285	
		(20.531)	(27.539)	
Gender of HH head (Female = 1)		-13.317	-188.162	
		(149.196)	(213.436)	
Access to credit (0/1)		-133.204	-207.246	
		(216.709)	(228.268)	
HH is not migrant (0/1)		-215.275	-525.959**	
		(181.016)	(264.090)	
Farm land under cashew nuts (acres)		27.296*	74.250***	
		(14.184)	(19.163)	

Family labor (days per year)	-0.230	-0.414
	(0.535)	(0.729)
Hired labor (days per year)	3.555	-1.099
	(2.240)	(2.956)
Application of herbicide/pesticide /fungicide (litres)	1.605*	1.562***
	(0.833)	(0.593)
Soil quality good or excellent (0/1)	139.462	141.874
	(99.165)	(129.392)
Soil steep (0/1)	-29.935	299.060
	(279.001)	(279.080)
Received agricultural training (0/1)	-221.901**	-155.613
	(110.533)	(132.999)
District term	13.038	28.434***
	(12.502)	(8.621)
Constant	1196.111**	1187.405**
	(476.054)	(506.291)
OME1		
HH size	41.675***	31.777
	(13.395)	(44.528)
Age of HH head	-6.455***	-6.057
	(1.322)	(4.995)
Years of education of HH head	-64.572***	-30.066
	(14.096)	(20.365)
Gender of HH head (Female = 1)	-220.908*	-150.474
	(128.547)	(195.885)
Access to credit (0/1)	-228.492	-246.183
	(219.421)	(615.170)
HH is not migrant (0/1)	-110.093	-162.840
	(212.551)	(401.280)
Farm land under cashew nuts (acres)	37.355***	57.523***
	(10.101)	(20.085)
Family labor (days per year)	1.211	1.337
	(0.856)	(1.315)
Hired labor (days per year)	0.725	2.163
	(2.055)	(3.661)
Application of	3.819***	3.576***

herbicide/pesticide /fungicide (litres)		
	(0.127)	(0.492)
Soil quality good or excellent (0/1)	192.155***	275.363
	(54.416)	(201.627)
Soil steep (0/1)	101.454	71.212
	(162.462)	(397.183)
Received agricultural training (0/1)	32.383	30.914
	(120.864)	(125.632)
District term	9.387	-18.728
	(11.973)	(13.907)
Constant	690.894*	565.397
	(380.566)	(572.815)
TME1		
HH size	0.038	0.038
	(0.038)	(0.041)
Age of HH head	-0.017***	-0.017***
	(0.006)	(0.006)
Years of education of HH head	-0.169***	-0.169***
	(0.027)	(0.030)
Gender of HH head (Female = 1)	-0.771***	-0.771***
	(0.178)	(0.255)
Access to credit (0/1)	-0.502*	-0.502*
	(0.259)	(0.265)
HH is not migrant (0/1)	-0.857**	-0.857***
	(0.369)	(0.277)
Farm land under cashew nuts (acres)	-0.021	-0.021*
	(0.013)	(0.012)
Family labor (days per year)	-0.001	-0.001
	(0.001)	(0.001)
Hired labor (days per year)	0.001	0.001
	(0.002)	(0.002)
Application of herbicide/pesticide /fungicide (litres)	0.000	0.000
	(0.001)	(0.000)
Soil quality good or excellent (0/1)	0.082	0.082
	(0.100)	(0.161)
Soil steep (0/1)	-0.030	-0.030

			(0.163)	(0.335)	
Received agricultural training (0/1)			0.695***	0.695***	
			(0.224)	(0.166)	
Constant			2.503***	2.503***	
			(0.773)	(0.574)	
Observations	339	339	339	339	339

Note: Standard errors in parentheses: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 3A.9: Effects of contract farming on food insecurity, full table with controls

	(1) OLS	(2) PSM-ATT	(3) IPWRA-ATE	(4) IPWRA-ATT	(5) IV
Contract farming	-0.637***	-0.705***	-0.692***	-0.770***	-0.577***
	(0.084)	(0.133)	(0.077)	(0.124)	(0.096)
HH size	-0.022				-0.026**
	(0.014)				(0.013)
Age of HH head	-0.003				-0.001
	(0.004)				(0.004)
Years of education of HH head	-0.035				-0.032
	(0.028)				(0.028)
Gender of HH head (Female = 1)	0.110				0.126
	(0.141)				(0.129)
Access to credit (0/1)	0.038				0.059
	(0.126)				(0.136)
HH is not migrant (0/1)	-0.078				-0.121
	(0.142)				(0.108)
Farm land under cashew nuts (acres)	-0.028**				-0.026**
	(0.013)				(0.013)
Family labor (days per year)	0.001				0.001
	(0.001)				(0.001)
Hired labor (days per year)	-0.004**				-0.004***
	(0.002)				(0.002)
Application of herbicide/pesticide /fungicide (litres)	-0.001				-0.001*
	(0.000)				(0.000)
Soil quality good or excellent (0/1)	-0.260**				-0.288***
	(0.097)				(0.093)
Soil steep (0/1)	-0.322				-0.370*
	(0.215)				(0.215)

Received agricultural training (0/1)	-0.280**			-0.293***
	(0.100)			(0.092)
District term	-0.071**			-0.068***
	(0.029)			(0.026)
Constant	4.048***			3.974***
	(0.453)			(0.453)
Potential outcome mean				
0. No		2.435***	2.522***	
		(0.089)	(0.148)	
OME0				
HH size		0.001	0.015	
		(0.024)	(0.043)	
Age of HH head		0.001	0.005	
		(0.006)	(0.008)	
Years of education of HH head		-0.058**	-0.070**	
		(0.027)	(0.033)	
Gender of HH head (Female = 1)		-0.112	-0.500*	
		(0.205)	(0.278)	
Access to credit (0/1)		0.054	0.157	
		(0.159)	(0.327)	
HH is not migrant (0/1)		0.128	0.329	
		(0.194)	(0.284)	
Farm land under cashew nuts (acres)		-0.035**	-0.059***	
		(0.015)	(0.016)	
Family labor (days per year)		0.002***	0.003**	
		(0.001)	(0.001)	
Hired labor (days per year)		-0.006***	-0.005**	
		(0.002)	(0.002)	
Application of herbicide/pesticide /fungicide (litres)		-0.000	-0.000	
		(0.000)	(0.000)	
Soil quality good or excellent (0/1)		-0.213**	0.048	
		(0.099)	(0.161)	
Soil steep (0/1)		-0.302	-0.984**	
		(0.274)	(0.407)	
Received agricultural training (0/1)		-0.269**	-0.385***	

	(0.129)	(0.127)
District term	0.000	-0.011
	(0.010)	(0.016)
Constant	3.324***	3.173***
	(0.488)	(0.772)
OME1		
HH size	-0.030	-0.014
	(0.019)	(0.009)
Age of HH head	-0.004	-0.005
	(0.003)	(0.003)
Years of education of HH head	0.004	-0.016
	(0.016)	(0.032)
Gender of HH head (Female = 1)	0.405***	0.587***
	(0.091)	(0.062)
Access to credit (0/1)	0.592***	0.575***
	(0.085)	(0.103)
HH is not migrant (0/1)	0.160	0.063
	(0.214)	(0.340)
Farm land under cashew nuts (acres)	-0.047***	-0.045**
	(0.009)	(0.021)
Family labor (days per year)	-0.001	-0.001
	(0.001)	(0.001)
Hired labor (days per year)	-0.001	-0.001
	(0.001)	(0.003)
Application of herbicide/pesticide /fungicide (litres)	-0.001***	-0.001***
	(0.000)	(0.000)
Soil quality good or excellent (0/1)	-0.149	-0.128
	(0.252)	(0.128)
Soil steep (0/1)	-0.898***	-0.653***
	(0.150)	(0.102)
Received agricultural training (0/1)	-0.083	-0.156***
	(0.063)	(0.031)
District term	-0.037***	-0.011**
	(0.012)	(0.005)
Constant	3.044***	2.998***
	(0.642)	(0.701)
TME1		
HH size	0.038	0.038

			(0.038)	(0.038)	
Age of HH head			-0.017***	-0.017***	
			(0.006)	(0.006)	
Years of education of HH head			-0.169***	-0.169***	
			(0.027)	(0.027)	
Gender of HH head (Female = 1)			-0.771***	-0.771***	
			(0.178)	(0.178)	
Access to credit (0/1)			-0.502*	-0.502*	
			(0.259)	(0.259)	
HH is not migrant (0/1)			-0.857**	-0.857**	
			(0.369)	(0.369)	
Farm land under cashew nuts (acres)			-0.021	-0.021	
			(0.013)	(0.013)	
Family labor (days per year)			-0.001	-0.001	
			(0.001)	(0.001)	
Hired labor (days per year)			0.001	0.001	
			(0.002)	(0.002)	
Application of herbicide/pesticide /fungicide (litres)			0.000	0.000	
			(0.001)	(0.001)	
Soil quality good or excellent (0/1)			0.082	0.082	
			(0.100)	(0.100)	
Soil steep (0/1)			-0.030	-0.030	
			(0.163)	(0.163)	
Received agricultural training (0/1)			0.695***	0.695***	
			(0.224)	(0.224)	
Constant			2.503***	2.503***	
			(0.773)	(0.773)	
Observations	339	339	339	339	339

Note: Standard errors in parentheses: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 3A.10: Placebo test: distance to nearest firm on outcome variables

	(1) RCN output	(2) Food insecurity (HFIAS)
Distance to the nearest AP firm (in Km)	-4.44	0.00
	(5.27)	(0.01)
HH size	54.50*	-0.03**
	(29.34)	(0.01)

Age of HH head	-12.16**	0.00
	(4.98)	(0.00)
Years of education of HH head	-21.49	-0.01
	(28.78)	(0.03)
Gender of HH head (Female = 1)	-56.25	0.24*
	(122.46)	(0.13)
Access to credit (0/1)	-53.52	0.11
	(249.71)	(0.16)
HH is not migrant (0/1)	7.56	-0.06
	(142.58)	(0.15)
Farm land under cashew nuts (acres)	14.16	-0.02*
	(9.73)	(0.01)
Family labor (days per year)	0.48	0.00
	(0.51)	(0.00)
Hired labor (days per year)	3.72	-0.00**
	(2.35)	(0.00)
Application of herbicide/pesticide/fungicide (litres)	2.80***	-0.00*
	(0.58)	(0.00)
Soil quality good or excellent (0/1)	162.16*	-0.30***
	(88.37)	(0.09)
Soil steep (0/1)	-179.86	-0.36
	(346.42)	(0.22)
Received agricultural training (0/1)	-25.97	-0.42***
	(111.37)	(0.09)
District term	12.68	-0.11
	(50.40)	(0.07)
Constant	997.60**	3.51***
	(383.55)	(0.65)
Observations	339	339

Note: Standard errors in parentheses: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. OLS model.

7.3 Appendix for Chapter 4

7.3.1 Tables

Table 4A.1: List of products classified by processing category

Agrifood product sub-category	SITC 3-digit product code
Unprocessed	001, 034, 035, 036, 041, 043, 044, 045, 054, 057, 071, 072, 073, 075, 081, 222
Lightly processed	011, 012, 016, 017, 025, 037, 042, 046, 047, 048, 056, 058, 061, 073, 223, 411, 421, 422
Highly processed	022, 023, 024, 059, 062, 091, 098, 111, 112, 431

Table 4A.2: List of trade digitalization policy variables

Policy variable	Mechanisms and components
Paperless trade	Automated Customs System E-Payment of Customs Duties and Fees Internet connection available to Customs and other trade control agencies Electronic Single Window System Electronic submission of Customs declarations Electronic application and issuance of import and export permit Electronic Submission of Sea/Air Cargo Manifests Electronic application and issuance of Preferential Certificate of Origin Electronic Application for Customs Refunds
Cross-border NTMs	Laws and regulations for electronic transactions Recognized certification authority Electronic exchange of Customs Declaration/Certificate of Origin Electronic exchange of Sanitary & Phytosanitary Certificate Paperless collection of payment from a documentary letter of credit
Transparency	Publication of existing import-export regulations on the internet Stakeholders' consultation on new draft regulations (prior to their finalization) Advance publication/notification of new trade-related regulations before their implementation Advance ruling on tariff classification and origin of imported goods Independent appeal mechanism
Formalities	Risk management Pre-arrival processing Post-clearance audits Separation of release from final determination of customs duties, taxes, fees and charges

Establishment and publication of average release times
Trade facilitation measures for authorized operators
Expedited shipments
Acceptance of copies of original supporting documents required for import, export or transit formalities

Table 4A.3: List of gravity and socio-cultural control variables of the gravity model

Variables	Source
GDP PPP (in current USD)	CEPII-BACI / UN Comtrade
Distance between pair of trading countries	CEPII-BACI
RTA type and coverage between pair	WTO
RTA Coverage	WTO
Contiguity and/or shared border	CEPII-BACI
Common official or primary language	CEPII-BACI
Language is spoken by at least 9% of the population	CEPII-BACI
Common colonizer post 1945	CEPII-BACI
Pair in colonial relationship post 1945	CEPII-BACI
Origin of the legal system	CEPII-BACI
Pair current or former hegemon	CEPII-BACI
Pair ever/currently in colonial or dependency relationship	CEPII-BACI
WTO membership	WTO
EU membership	CEPII-BACI

Table 4A.4: Descriptive statistics of the control and gravity variables

	Mean (SD)
GDP PPP, in current thousands USD	4.25e+08 (1.78e+09)
Distance (in Km) between pair of trading countries	8442.60 (4703.00)
Dummy if there is an RTA between pair of trading countries	0.145 (0.35)
Contiguity and/or shared border	0.012 (0.11)
Common official or primary language	0.176 (0.38)
Language is spoken by at least 9% of the population	0.169 (0.37)
Common colonizer post 1945	0.119 (0.32)
Pair in colonial relationship post 1945	0.006 (0.08)

Pair ever in colonial or dependency relationship	0.008
	(0.09)
Pair currently in colonial or dependency relationship	0.001
	(0.04)
Origin of the legal system	2.043
	(1.06)
WTO membership	0.676
	(0.47)
EU membership	0.117
	(0.32)
Observations	8,217,072

Table 4A.5: Correlation matrix of the policy variables

	PT	**CB**	**Transparency**	**Formalities**	**Institutions**
PT	1.000				
CB	0.785***	1.000			
Transparency	0.705***	0.661***	1.000		
Formalities	0.800***	0.693***	0.720***	1.000	
Institutions	0.453***	0.501***	0.518***	0.501***	1.000

Note: Results are symmetrical between origin and destination. ***: significant at the 1% level with Bonferroni-corrected p-value.

Table 4A.6: Effects of trade transparency digitalization on agrifood trade: All food groups

Log of Agrifood Trade	(1)	(2)	(3)	(4)	(5)	(6)
Transparency, importers	0.395***		0.310***		0.313***	
	(0.033)		(0.027)		(0.024)	
Transparency, exporters		0.536***		0.416***		0.430***
		(0.035)		(0.029)		(0.027)
Constant	-8.483***	-8.451***	-8.713***	-8.376***	-2.065***	-6.079***
	(0.779)	(0.804)	(0.207)	(0.206)	(0.668)	(0.536)
Time trend	YES	YES	YES	YES	YES	YES
Gravity variables	YES	YES	YES	YES	YES	YES
Socioeconomic Controls	NO	NO	YES	YES	YES	YES
Fixed Effects	NO	NO	NO	NO	YES	YES
Observations	2,618,182	2,618,182	2,618,182	2,618,182	2,618,182	2,618,182
Importing countries	139	180	139	180	139	180
Exporting countries	180	139	180	139	180	139
Pseudo R^2	0.405	0.370	0.407	0.377	0.416	0.381
AIC	2.01e+7	1.42e+7	1.89e+7	1.25e+7	1.71e+7	1.14e+7

Note: Standard errors in parentheses and significance levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Fixed-effects models: R^2 captures the Pseudo-R^2. Policy variables: measured in log.

Table 4A.7: Effects of trade formalities digitalization on agrifood trade: All food groups

Log of Agrifood Trade	(1)	(2)	(3)	(4)	(5)	(6)
Formalities, importers	0.310***		0.245***		0.261***	
	(0.029)		(0.029)		(0.026)	
Formalities, exporters		0.635***		0.530***		0.532***
		(0.029)		(0.032)		(0.030)
Constant	-8.862***	-8.176***	-9.114***	-8.088***	-2.179***	-5.704***
	(0.820)	(0.840)	(0.202)	(0.203)	(0.669)	(0.535)
Time trend	YES	YES	YES	YES	YES	YES
Gravity variables	YES	YES	YES	YES	YES	YES
Socioeconomic Controls	NO	NO	YES	YES	YES	YES
Fixed Effects	NO	NO	NO	NO	YES	YES
Observations	2,618,182	2,618,182	2,618,182	2,618,182	2,618,182	2,618,182
Importing countries	139	180	139	180	139	180
Exporting countries	180	139	180	139	180	139
Pseudo R^2	0.397	0.361	0.401	0.373	0.415	0.382
AIC	1.64e+7	1.77e+7	1.47e+7	1.58e+7	1.35e+7	1.50e+7

Note: Standard errors in parentheses and significance levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Fixed-effects models: R^2 captures the Pseudo-R^2. Policy variables: measured in log.

Table 4A.8: Effects of paperless trade digitalization on agrifood trade in SSA: All food groups

Log of Agrifood Trade	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
PT, importers	0.246**		0.262***		0.166***		0.269***		0.208***		0.167***	
	(0.041)		(0.048)		(0.059)		(0.041)		(0.041)		(0.055)	
PT, exporters		0.642***		0.475***		0.695***		0.449***		0.499***		0.654***
		(0.052)		(0.053)		(0.063)		(0.047)		(0.055)		(0.052)
CB, importers									0.056***			
									(0.020)			
CB, exporters										0.025		
										(0.016)		
Transparency, destination					0.042						0.047	
					(0.099)						(0.085)	
Transparency, origin						0.497***						0.413***
						(0.070)						(0.056)
Formalities, destination					0.263**						0.271***	
					(0.105)						(0.099)	
Formalities, origin						1.575***						1.350***
						(0.124)						(0.098)
Constant	-9.323***	-10.220***	-11.180***	-11.580***	-10.790***	-14.288***	-8.501***	-4.210	-8.338***	-3.892	-7.866***	-6.240**
	(0.590)	(1.096)	(0.599)	(0.721)	(0.681)	(0.721)	(1.289)	(2.579)	(1.295)	(2.661)	(1.331)	(2.544)
Time trend	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Gravity variables	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Socioeconomic Controls	NO	NO	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Fixed Effects	NO	NO	NO	NO	NO	NO	YES	YES	YES	YES	YES	YES
Observations	714,748	500,894	680,938	481,344	680,938	481,344	620,402	437,828	620,402	437,828	620,402	437,828
Importing countries	32	139	32	139	32	139	32	139	32	139	32	139
Exporting countries	139	32	139	32	139	32	139	32	139	32	139	32
Pseudo R^2	0.058	0.042	0.100	0.095	0.100	0.098	0.300	0.286	0.301	0.286	0.301	0.291
AIC	2.02e+6	1.42e+6	1.88e+6	1.23e+6	1.88e+6	1.22e+6	1.71e+6	1.14e+6	1.57e+6	8.84e+5	1.71e+6	1.13e+6

Note: Standard errors in parentheses and significance levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Fixed-effects models: R^2 captures the Pseudo-R^2. Policy variables: measured in log.

Table 4A.9: Effects of cross-border NTMs digitalization on agrifood trade in SSA: All food groups

Log of Agrifood Trade	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
CB, importers	0.094***		0.076***		0.054**		0.089***		0.066***	
	(0.023)		(0.022)		(0.022)		(0.019)		(0.019)	
CB, exporters		0.098***		0.035*		0.068***		0.042***		0.063***
		(0.020)		(0.019)		(0.020)		(0.014)		(0.016)
Transparency, destination					0.032				0.026	
					(0.099)				(0.083)	
Transparency, origin						0.654***				0.544***
						(0.069)				(0.057)
Formalities, destination					0.395***				0.396***	
					(0.094)				(0.085)	
Formalities, origin						1.330***				1.091***
						(0.134)				(0.108)
Constant	-9.553***	-11.410***	-11.960***	-13.180***	-10.641***	-15.426***	-9.088***	-7.274***	-7.781***	-9.915***
	(0.590)	(1.071)	(0.551)	(0.701)	(0.682)	(0.743)	(1.301)	(2.589)	(1.349)	(2.666)
Time trend	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Gravity variables	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Socioeconomic Controls	NO	NO	YES	YES	YES	YES	YES	YES	YES	YES
Fixed Effects	NO	NO	NO	NO	NO	NO	YES	YES	YES	YES
Observations	714,748	500,894	680,938	481,344	680,938	481,344	620,402	437,828	620,402	437,828
Importing countries	32	139	32	139	32	139	32	139	32	139
Exporting countries	139	32	139	32	139	32	139	32	139	32
Pseudo R^2	0.058	0.040	0.100	0.094	0.100	0.096	0.291	0.286	0.301	0.301
AIC	1.84e+6	1.13e+6	1.87e+6	1.24e+6	2.83e+6	4.09e+6	1.57e+6	8.90e+5	1.71e+6	1.14e+6

Note: Standard errors in parentheses and significance levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Fixed-effects models: R^2 captures the Pseudo-R^2. Policy variables: measured in log.

Table 4A.10: Effects of paperless trade digitalization on agrifood trade in Asia: All food groups

Log of Agrifood Trade	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
PT, importers	0.048		0.040		0.043		0.058**		0.037		0.034	
	(0.030)		(0.029)		(0.032)		(0.025)		(0.026)		(0.027)	
PT, exporters		0.627***		0.584***		0.368***		0.584***		0.260***		0.376***
		(0.048)		(0.045)		(0.050)		(0.043)		(0.040)		(0.047)
CB, importers									0.028**			
									(0.011)			
CB, exporters										0.331***		
										(0.019)		
Transparency, destination					0.200***						0.213***	
					(0.058)						(0.049)	
Transparency, origin						0.272***						0.221***
						(0.100)						(0.091)
Formalities, destination					0.183***						0.214***	
					(0.066)						(0.057)	
Formalities, origin						0.791***						0.843***
						(0.087)						(0.085)
Constant	-9.661***	-7.318***	-10.280***	-8.132***	-9.113***	-6.748***	-3.646*	-5.643***	-3.400*	-3.869***	-2.541	-3.383***
	(0.900)	(0.779)	(0.494)	(0.509)	(0.516)	(0.527)	(2.066)	(0.850)	(2.031)	(0.848)	(2.062)	(0.850)
Time trend	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Gravity variables	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Socioeconomic Controls	NO	NO	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Fixed Effects	NO	NO	NO	NO	NO	NO	YES	YES	YES	YES	YES	YES
Observations	620,080	878,692	539,028	783,380	539,028	783,380	524,032	747,316	524,032	747,316	524,032	747,316
Importing countries	30	139	30	139	30	139	30	139	30	139	30	139
Exporting countries	139	30	139	30	139	30	139	30	139	30	139	30
Pseudo R^2	0.236	0.250	0.279	0.282	0.282	0.282	0.415	0.395	0.415	0.403	0.417	0.399
AIC	3.56e+6	4.81e+6	3.25e+6	4.42e+6	3.25e+6	4.40e+6	2.89e+06	4.12e+6	2.51e+6	3.79e+6	2.88e+6	4.09e+6

Note: Standard errors in parentheses and significance levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Fixed-effects models: R^2 captures the Pseudo-R^2. Policy variables in log.

Table 4A.11: Effects of cross-border NTMs digitalization on agrifood trade in Asia: All food groups

Log of Agrifood Trade	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
CB, importers	0.019		0.023*		0.021		0.034***		0.009	
	(0.014)		(0.013)		(0.013)		(0.010)		(0.010)	
CB, exporters		0.310***		0.345***		0.279***		0.381***		0.308***
		(0.017)		(0.018)		(0.017)		(0.019)		(0.017)
Transparency, destination					0.181***				0.192***	
					(0.059)				(0.050)	
Transparency, origin						0.183***				0.099
						(0.092)				(0.087)
Formalities, destination					0.155***				0.186***	
					(0.057)				(0.048)	
Formalities, origin						0.524***				0.610***
						(0.086)				(0.086)
Constant	-9.705***	-7.264***	-10.250***	-7.672***	-9.244***	-6.788***	-3.563*	-4.789***	-2.391	-2.867***
	(0.906)	(0.775)	(0.497)	(0.468)	(0.517)	(0.519)	(2.076)	(0.817)	(2.065)	(0.855)
Time trend	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Gravity variables	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Socioeconomic Controls	NO	NO	YES	YES	YES	YES	YES	YES	YES	YES
Fixed Effects	NO	NO	NO	NO	NO	NO	YES	YES	YES	YES
Observations	620,080	878,692	539,028	783,380	539,028	783,380	524,032	747,316	524,032	747,316
Importing countries	30	139	30	139	30	139	30	139	30	139
Exporting countries	139	30	139	30	139	30	139	30	139	30
Pseudo R^2	0.236	0.250	0.280	0.287	0.282	0.287	0.415	0.402	0.417	0.404
AIC	3.11e+6	4.50e+6	3.25e+6	4.38e+6	3.25e+5	4.37e+5	2.89e+6	4.07e+6	2.88e+6	4.06e+6

Note: Standard errors in parentheses and significance levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Fixed-effects models: R^2 captures the Pseudo-R^2. Policy variables: measured in log.