Agriculture-Nutrition Linkages in the Kenyan Small Farm Sector: The Role of Commercialization, Technology Adoption, and Extension

Dissertation
to obtain the Ph.D. degree
in the International Ph. D. Program for Agricultural Sciences in Goettingen (IPAG)
at the Faculty of Agricultural Sciences,
Georg-August-University Göttingen, Germany

presented by

Sylvester Ochieng Ogutu

born in Homa Bay, Kenya

Göttingen, March 2018

D7

1. Name of supervisor: Prof. Dr. Matin Qaim
2. Name of co-supervisor: Prof. Dr. Meike Wollni
3. Member of examination committee: Prof. Dr. Ludwig Theuvsen

Date of dissertation: 17th May 2018

Summary

Global hunger, malnutrition, and poverty have declined over the past decades, but a situation of food and nutrition security for all is yet to be achieved, and extreme poverty is still widespread. Around 800 million people globally are still chronically undernourished, about 2 billion people suffer from micronutrient deficiencies which pose serious health consequences, and 767 million people live in extreme poverty – below 1.90 US dollars a day. Hence, eradication of hunger, malnutrition, and extreme poverty continue to rank high on the development policy agenda. A large proportion of the people affected are smallholder farmers in developing countries who depend on agriculture as a source of food and income. Therefore, a key question for improving nutrition is how to make smallholder agriculture more nutrition-sensitive. Previous studies proposed the use of agricultural technologies specifically designed to improve nutrition (e.g. biofortified crop technologies), higher on-farm diversification, and strengthening of smallholder access to markets as part of key strategies to improve nutrition in the small farm sector. These strategies could also reduce poverty in farm households, for instance, through increased labor and land productivity, and cash income gains from market sales.

This dissertation contributes to the emerging research on agriculture-nutrition linkages, particularly addressing three research gaps. The first research gap relates to the adoption of biofortified crops and other pro-nutrition innovations. One particular challenge with the adoption of biofortified technologies is a low incentive among farmers to adopt such technologies due to limited awareness of their nutritional benefits. Agricultural extension could play a crucial role in creating the required awareness to enhance adoption. However, previous research on how to improve extension services primarily focused on the dissemination of technologies with potential agronomic benefits, not on pro-nutrition technologies. The second research gap relates to the link between agricultural commercialization and nutrition. While it is often assumed that subsistence production is particularly important for smallholder diets, effects of agricultural commercialization on household nutrition and dietary quality have hardly been analyzed. The third research gap relates to wider welfare effects of agricultural commercialization. Previous studies evaluated the impacts of commercialization on smallholder income and poverty, but potential effects on other livelihood dimensions and multidimensional poverty are much less understood. These research gaps are addressed with primary data collected from approximately 800 smallholder

farm households in Western Kenya and the use of various econometric and experimental techniques.

This dissertation comprises three essays that contribute to the literature on agriculture and nutrition linkages in different ways. In the first essay, we conduct a randomized controlled trial (RCT) to evaluate the effects of new agricultural extension designs on the adoption of a pro-nutrition technology. In particular, we examine how agricultural training can be combined with training in nutrition and marketing to increase farmers' adoption of a new bean variety biofortified with iron and zinc. To the best of our knowledge, this is the first study to analyze how improved designs of agricultural extension can contribute to making smallholder farming more nutrition-sensitive. This analysis is based on panel data from two survey rounds. Difference-in-difference estimates show that intensive agricultural training tailored to local conditions can increase technology adoption considerably. In less than one year, adoption of biofortified beans increased from almost zero to more than 20%. With additional nutrition training, adoption further increased by 10-12 percentage points, since this helped farmers to better appreciate the technology's nutritional benefits. These results suggest that effective nutrition training through agricultural extension services is possible. Providing additional marketing training did not lead to further adoption effects, possibly because the study period may have been too short to measure these effects properly. Based on the results we argue that closer cooperation between agricultural extension and nutrition and health organizations can be an effective way to promote pro-nutrition innovations among smallholder farm households.

In the second essay, we evaluate the effects of commercialization on household food security and dietary quality, with a particular focus on calorie and micronutrient consumption. The few previous studies on the effects of commercialization on nutrition examined impacts in terms of calorie intake and child anthropometrics, but not dietary quality as we do. We further depart from previous studies by estimating average and continuous treatment effects, and also examine possible transmission channels through which commercialization affects farm household nutrition by looking at the role of income, gender, and possible substitution between the consumption of own-produced (subsistence) and purchased foods. This analysis builds on cross-sectional survey data from 805 farm households. We use a control function approach to address potential endogeneity problems. Generalized propensity scores are employed to estimate continuous treatment effects. The results show that commercialization significantly improves food security and dietary quality measured in terms of calorie, zinc, and iron consumption. For vitamin A, overall effects are statistically insignificant.

Commercialization contributes to higher incomes and additional nutrient intake from purchased foods, but it does not reduce the consumption of nutrients from own-produced foods, even after controlling for farm size, possibly due to higher productivity on more commercialized farms. We conclude that enhancing market access is important not only for rural economic growth, but also for making smallholder agriculture more nutrition-sensitive.

In the third essay, we examine the impacts of commercialization on multidimensional poverty. Previous studies examined the effects of commercialization on poverty with income-based poverty indicators, but income-poverty indicators are prone to measurement error and do not capture the many dimensions of poverty (welfare), such as education, health, and living standard. We use cross-sectional data from a sample of 805 farm households and various econometric models to measure average and heterogeneous treatment effects of commercialization on multidimensional poverty. A 10 percentage point increase in the degree of commercialization reduces the likelihood of being multidimensionally poor by 2.3 percentage points and the multidimensional poverty index (intensity) – share of total deprivations in years of schooling, child school attendance, calorie consumption, dietary quality, sanitation, access to safe drinking water, electricity, cooking fuel, floor material and asset ownership – by 1.5 percentage points. Quantile regressions show that commercialization significantly reduces multidimensional poverty across the different quantiles. Commercialization also reduces income poverty. The absolute gains in per capita income are larger for richer households. We conclude that commercialization can contribute to multidimensional and income poverty reduction in smallholder farm households, but may possibly also lead to higher income inequality.

We draw several conclusions and policy implications from the three essays in this dissertation. From the first essay, we conclude that combining agricultural and nutrition training in agricultural extension approaches is feasible, and can contribute to making smallholder farming more nutrition-sensitive. Therefore, closer cooperation between agricultural extension and nutrition and health organizations could provide a cost-effective way to promote pro-nutrition innovations in the small farm sector. However, further research is needed on how the design of agricultural extension approaches can be improved in order to increase the adoption of pro-nutrition technologies. Our study is only an initial step in this direction. From the second essay, we conclude that commercialization can contribute significantly to improved nutrition in the small farm sector. The important policy implication is that enhancing market access is a key strategy to make smallholder agriculture more nutrition-sensitive. But we stress that commercialization alone will not suffice to address all types of malnutrition. Commercialization increases cash income, but the consumption of certain micronutrients – such as vitamin A – does not seem to be particularly responsive to income growth. Hence, more specific complementary interventions may be needed. From the third essay, we conclude that commercialization contributes to multidimensional poverty reduction in the small farm sector. Hence, smallholder access to markets should be strengthened. Further research is needed to quantify the long-term effects of commercialization on multidimensional poverty and to provide more evidence in different contexts.

Acknowledgements

Foremost, I thank the Almighty God for giving me strength, good health, and the ability to successfully complete my Ph.D. studies. I also express my sincere gratitude to my supervisor Prof. Dr. Matin Qaim for offering me a scholarship for doctoral studies in his research group, and for his invaluable academic guidance throughout my studies. His feedbacks were insightful and timely, making the burden that Ph.D. studies can be much lighter. Indeed I could not wish for a better mentor. My special thanks also go to Prof. Dr. Meike Wollni and Prof. Dr. Ludwig Theuvsen for co-supervising and serving on my dissertation examination committee.

I thank all my colleagues at the Chair of International Food Economics and Rural Development, and GlobalFood Research Training Group (RTG 1666) at the University of Goettingen for providing a conducive academic environment. I am very grateful to Ilona Hoepfner for her administrative assistance and Dr. Betty Debela for providing interesting comments on my essays. I express my special thanks to Dr. Kibrom Sibhatu, Dr. Priscilla Wainaina, Dr. Dennis Ochieng, Dr. Kenneth Sibiko, Dr. Haruna Sekabira, Zewdu, Rosina and Makaiko for their friendship and for making my stay in Goettingen more enjoyable. I often recollect the summer trips and the get-togethers that we had with great pleasure.

This research was conducted in collaboration with Africa Harvest Biotech Foundation International and colleagues in the Department of Agricultural Economics at the University of Nairobi. Hence, I am grateful for the support and research cooperation received from both organizations. I also express my sincere appreciation to our team of enumerators in both baseline and follow-up surveys, who in spite of the harsh fieldwork conditions were able to collect good quality data. I gratefully acknowledge the financial support for this research which came from the German Federal Ministry of Food and Agriculture (BMEL) based on the decision of the Parliament of the Federal Republic of Germany (grant number 2813FSNu01).

Last but not least, I wish to express my deepest gratitude to my dear wife Dorothy, for her unwavering support and encouragement throughout my Ph.D. studies, and for taking good care of our beloved son Ethan in my long absence. I also express my special thanks to my mother Pamela Ogutu, for her steadfast love, prayers, encouragement, and for teaching me virtue, to my father Julius Ogutu for teaching me the value of education and hard work, to my siblings Anne and Loice, for their moral support and love. This dissertation is dedicated to everyone in my family – you are special!

Table of Contents

List of Tables

List of Figures

List of Abbreviations

AE	Adult Male Equivalent
CF	Control Function
FAO	Food and Agriculture Organization
GDP	Gross Domestic Product
GLM	Generalized Linear Model
GPS	Generalized Propensity Score
HDDS	Household Dietary Diversity Score
IFPRI	International Food Policy Research Institute
ITT	Intent-To-Treat
IV	Instrumental Variable
IZiNCG	International Zinc Nutrition Consultative Group
KALRO	Kenya Agricultural Livestock Research Organization
KNBS	Kenya National Bureau of Statistics
LPM	Linear Probability Model
MDDW	Minimum Dietary Diversity for Women
MPI	Multidimensional Poverty Index
OLS	Ordinary Least Squares
OPHI	Oxford Poverty and Human Development Initiative
PPP	Purchasing Power Parity
RCT	Randomized Controlled Trial
TLU	Tropical Livestock Unit
TOT	Treatment-On-The-Treated
UNU	United Nations University
US	United States
USAID	United States Agency for International Development
WHO	World Health Organization

1 General Introduction

1.1 Background

In spite of the decline in global hunger, malnutrition, and poverty over the past decades, food and nutrition insecurity and extreme poverty still persist. Approximately 800 million people globally are chronically undernourished (energy/calorie deficient), and about 2 billion people suffer from micronutrient deficiencies, which pose serious negative health consequences, such as vulnerability to communicable diseases, impaired physical and cognitive human development, and pre-mature deaths (IFPRI 2017; FAO 2017). The micronutrient deficiencies are mostly in intakes of vitamins and minerals, such as vitamin A, iron, and zinc. The economic implications of malnutrition amount to 11% loss of gross domestic product (GDP) annually in Africa and Asia, but preventing malnutrition delivers 16 US dollars in returns on investment for every US dollar spent (IFPRI 2016). About 767 million people worldwide also live in extreme poverty – below 1.90 US dollars a day (World Bank 2016). Hence, eradication of hunger, malnutrition, and extreme poverty continue to rank high on the development policy agenda.

The majority of the people affected by hunger, malnutrition, and poverty are smallholder farmers in developing countries – especially in sub-Saharan Africa and south Asia – who depend on agriculture for their livelihoods (World Bank 2007). Hence, agricultural development provides an important pathway to improved nutrition and poverty reduction. For instance, growth in agricultural GDP has been shown to be 2-4 times more effective in reducing poverty than growth in other sectors of the economy (World Bank 2015; de Janvry and Sadoulet 2009). But the principal question for improving nutrition is how to make smallholder farming more nutrition-sensitive. “Nutrition-sensitive agriculture is a food-based approach to agricultural development that puts nutritionally rich foods, dietary diversity, and food fortification at the heart of overcoming malnutrition and micronutrient deficiencies” (FAO 2014). It seeks to maximize agriculture’s contribution to nutrition. Improved nutrition is crucial for growth in health, education, employment, female empowerment, and poverty and inequality reduction that also affect nutrition outcomes (IFPRI 2017).

The question of how smallholder agriculture can be made more nutrition-sensitive has recently been analyzed by studies which can be classified in three strands. The first strand of literature has focused on on-farm diversification strategies, where a growing body of literature has examined the relationship between on-farm production diversity and dietary diversity

(Jones, Shrinivas, and Bezner-Kerr 2014; Sibhatu, Krishna, and Qaim 2015; Koppmair, Kassie, and Qaim 2017; Jones 2017). Most of these studies conclude that there is a positive association between on-farm production diversity and dietary diversity. But the relationship can be complex since the strength of the associations may vary depending on differences in gender roles within households, household wealth status, the relative degree of a household's market orientation, and the nature of farm diversity (Jones et al. 2014). Furthermore, when production diversity is already high, the associations may not be significant or may even turn negative due to foregone income benefits of on-farm specialization (Sibhatu et al. 2015). This suggests that increasing on-farm production diversity may not always be the most effective way to improve dietary diversity in smallholder households.

The second strand of literature has focused on promotion or dissemination of pro-nutrition technologies (Gilligan 2012; Jones and de Brauw 2015; Bouis and Saltzman 2017). Common examples of such pro-nutrition technologies are biofortified staple crops, which were bred to contain higher amounts of micronutrients, such as orange-fleshed sweet potatoes enhanced with provitamin A or high-iron rice and wheat (Bouis and Saltzman 2017; Jones and de Brauw 2015). Biofortification of food staples has several advantages. Staples are consumed daily, making biofortification pro-poor, cost-effective once adopted, and sustainable since biofortified crops can be replanted for years. However, success with biofortification has been limited by low adoption of the technologies and the lack of more cost-effective delivery strategies (Gilligan 2012). Hence, developing effective strategies for encouraging adoption and diffusion of biofortified crop technologies is crucial for improving nutrition in smallholder farm households.

The third strand of literature has explored the relationship between smallholders' access to markets and diets. This body of literature is very thin. The few studies that examined the link between market access and diets conclude that a positive relationship exists between smallholder access to markets and dietary diversity (Sibhatu et al. 2015; Koppmair et al. 2017). These studies also find that, access to market has a stronger positive association with dietary diversity, compared to on-farm diversity. However, the linkages between commercialization – measured as share of total farm output sold – and farm household dietary quality are much less understood. Commercialization may influence nutrition through various channels, such as changes in income, availability of own-produced foods, and differences in gender roles within the farm household (von Braun and Kennedy 1994; Carletto et al. 2015). Income gains can increase economic access to food, but a substitution of purchased food for own-produced food may also affect dietary quality, possibly increasing the consumption of calories but not necessarily micronutrients (Popkin, Adair, and Ng 2012; Remans et al. 2015). Changes in gender roles may occur because men often take stronger control of farm production and income during the process of commercialization (von Braun and Kennedy 1994), and male-controlled income is often spent less on dietary quality and nutrition than female-controlled income (Fischer and Qaim 2012a). Hence, understanding how these mechanisms can affect dietary quality is important for making agriculture more nutrition-sensitive. Beyond income and nutrition effects, commercialization may also have implications on multidimensional household poverty – deprivation in education, health, and living standards – which has rarely been examined. This dissertation contributes to the second and third strands of literature. We justify our contribution to the literature below.

1.2 Problem Statement and Objectives

Promotion of biofortified crops is one of the recent strategies that have been employed to eradicate hunger and micronutrient malnutrition. However, success with biofortification has been limited by low adoption of the technologies due to limited appreciation of their nutritional benefits by farmers, and lack of more cost-effective delivery strategies tailored to local conditions (Gilligan 2012). Recent research showed that the adoption of pro-nutrition technologies is higher in settings where farmers have a good understanding of the technologies' agronomic and nutritional attributes (de Brauw, Eozenou, and Moursi 2015; de Brauw et al. 2015; de Groote et al. 2016). This implies that agricultural extension could play a crucial role in technology dissemination. Indeed, agricultural extension services are designed to facilitate technology transfer and improve innovation processes in the farming sector, but concrete evidence with pro-nutrition technologies hardly exists. More generally, the effects of agricultural extension on adoption of innovations is rather mixed (Anderson and Feder 2004; Goodhue, Klonsky, and Mohapatra 2010; Läpple and Hennesy 2015). Hence, improvement in the design of agricultural extension is needed. While previous studies have analyzed how agricultural training components could be improved to increase farmers' adoption of agronomic innovations (Davis et al. 2012; Läpple and Hennessy 2015), research that has developed and tested new extension approaches for the effective dissemination of pro-nutrition technologies hardly exists. Here, we address this research gap with a randomized controlled trial (RCT) in Kenya. In particular, we evaluate how agricultural training can be combined with training in nutrition and marketing to increase farmers' adoption of a new bean variety biofortified with iron and zinc.

The second research gap that we identified contributes to the strand of literature on market access. Recent research showed that smallholder access to markets can improve nutrition – dietary diversity – in the small farm sector (Sibhatu et al. 2015; Koppmair et al. 2017). However, these studies captured farmers' access to markets only in terms of simple proxies such as market distance. Moreover, the dietary indicators used have shortcomings: most studies used household dietary diversity scores, which are suitable for measuring household food security, but not dietary quality (Kennedy, Ballard, and Dop 2013). Other studies that have measured market access more precisely in terms of the level of commercialization – share of total farm output sold – have analyzed impacts on household welfare, mostly in terms of income or assets ownership (Tipraqsa and Schreinemachers 2009, Carletto, Kilic, and Kirk 2011; Muriithi and Matz 2015), but not nutrition. The few studies that have explicitly analyzed the effects of commercialization on nutrition, looked at nutrition primarily in terms of calorie consumption and child anthropometrics (von Braun and Kennedy 1994; Carletto, Corral, and Guelfi 2017), not dietary quality. We add to this existing literature by analyzing the effects of commercialization on food security and dietary quality, measured in terms of calorie and micronutrient consumption at the household level. We further depart from previous studies by estimating average and continuous treatment effects, and also analyze possible transmission channels through which commercialization influences farm household nutrition by examining the role of income, gender, and possible substitution between the consumption of own-produced and purchased foods.

The third research gap relates to the strand of literature on the implications of commercialization on household poverty. While previous studies have analyzed effects of commercialization on income or asset ownership, the implications on farm household multidimensional poverty have received much less attention. Previous studies that examined the effects of commercialization on poverty used income-based poverty indicators (Olwande and Smale 2014; Muriithi and Matz 2015; Muricho et al. 2017). Although income poverty indicators are important and widely used, they are prone to measurement error and do not capture the many dimensions of poverty (welfare), such as education, health, and living standards. Consequently, some studies have stressed the need to measure household welfare using alternative or additional indicators (Muriithi and Matz 2015; Alkire and Santos 2014). We contribute to the literature by providing new insights on the effects of commercialization on multidimensional poverty, measured in three dimensions namely deprivation in household education (household head education and child school attendance), health (household calorie intake and dietary quality), and living standard (sanitation, access to safe drinking water, electricity, cooking fuel, floor material and asset ownership).

Following the identified research gaps, we aim to contribute to a better understanding of agriculture-nutrition linkages by addressing the following specific research objectives:

1. To evaluate the impact of different extension designs – stand-alone agricultural extension, agricultural extension combined with nutrition training, and agricultural extension combined with nutrition and market access training – on the adoption of a biofortified crop (new bean variety biofortified with iron and zinc).
2. To evaluate average and continuous treatment effects of commercialization on household food security and dietary quality, measured in terms of calorie and micronutrient consumption, and to analyze the possible channels through which commercialization affects nutrition, especially focusing on income, gender roles, and possible substitution between own-produced and purchased foods in household diets.
3. To examine average and heterogeneous treatment effects of commercialization on farm household income and multidimensional poverty.

As mentioned, this dissertation includes three essays, each one addressing one of these three objectives. Our research findings will provide useful insights to policymakers on how agricultural extension and commercialization can contribute to making agriculture more nutrition-sensitive, and how commercialization can contribute to the reduction of multidimensional poverty in the small farm sector.

1.3 Data and Study Context

This study is based on data from a sample of smallholder farm households in rural parts of Western Kenya. In Kenya, smallholder farming accounts for 75% of total agricultural output (Olwande et al. 2015). Many of these smallholders are also poor. About half of the Kenyan population is income poor, with 51% and 33% of the rural and urban populations living below the poverty line, respectively (Wiesmann et al. 2016). Approximately three-quarters of the Kenyan population reside in rural areas and depend on agriculture for their livelihood (World

Bank 2017). Like most other countries of sub-Saharan Africa, malnutrition problems are widespread in the Kenyan small farm sector (Muthayya et al. 2013; KNBS 2015). The prevalence of undernourishment is 19%, of child stunting is 26%, while the prevalence of anaemia in women of reproductive age is 27% (FAO 2017). The data used in this dissertation were collected between October and December 2015, and in the same months the following year. The first essay uses two rounds of panel data, while the second and third essays are based on the cross-sectional data collected in 2015. Further details on the sampling procedure and data used are provided in the essays. The survey questionnaire is included in appendix B1.

1.4 Description of Co-authors' Contributions

This dissertation, and the three essays in particular, benefited from the contributions of various people. I, Sylvester Ochieng Ogutu (S.O.O.), am the first and the main author of all the three essays. In the following I mention the co-authors and their specific roles:

- Essay 1 (chapter 2 of this dissertation) is co-authored with Andrea Fongar, Theda Gödecke, Lisa Jäckering, Henry Mwololo, Michael Njuguna, Meike Wollni and Matin Qaim. All authors jointly designed the research and the RCT. S.O.O., A.F., L.J., and H.W. collected the data. S.O.O. analyzed and interpreted the data and wrote the first draft of the paper. All authors commented on the draft paper and approved the final version.
- Essay 2 (chapter 3 of this dissertation) is co-authored with Theda Gödecke and Matin Qaim. S.O.O. developed the research idea, collected the data, analyzed and interpreted the data, and wrote the essay. T.G. and M.Q. commented on the essay at various stages and approved the final version.
- Essay 3 (chapter 4 of this dissertation) is co-authored with Matin Qaim. S.O.O. developed the research idea, collected the data, analyzed and interpreted the data and wrote the essay. M.Q. commented on the essay at various stages.

1.5 Dissertation Outline

The remainder of this dissertation is organized as follows. Chapter 2 presents the first essay, which evaluates the impact of three different extension models – stand-alone agricultural extension, agricultural extension combined with nutrition training, and agricultural extension combined with nutrition and market access trainings – on the adoption of a new bean variety biofortified with iron and zinc, using an RCT approach. Chapter 3 contains the second essay, which estimates average and continuous treatment effects of commercialization on household food security and dietary quality, and also analyzes the possible channels through which commercialization affects nutrition by examining the role of income, gender, and possible substitution between the consumption of own-produced and purchased foods. Chapter 4 presents the third essay, which examines average and heterogeneous treatment effects of commercialization on farm household multidimensional and income poverty. Chapter 5 contains the overall conclusions and discusses policy implications and study limitations.

2 How to Make Farming and Agricultural Extension More Nutrition-Sensitive: Evidence from a Randomized Controlled Trial in Kenya

Abstract

We analyze how agricultural extension can be made more effective in terms of increasing smallholder farmers' adoption of pro-nutrition technologies, such as biofortified crops. In a randomized controlled trial with farmers in Western Kenya, we implemented several extension treatments and evaluated their effects on the adoption of beans that were biofortified with iron and zinc. Difference-in-difference estimates show that intensive agricultural training tailored to local conditions can increase technology adoption considerably. Within less than one year, adoption of biofortified beans increased from almost zero to more than 20%. Providing additional nutrition training further increased adoption by 10-12 percentage points, as this has helped farmers to better appreciate the technology's nutritional benefits. These results suggest that effective nutrition training through agricultural extension services is possible. Providing marketing training did not lead to additional adoption effects, although the study period may have been too short to measure these effects properly. This study is a first attempt to analyze how improved designs of agricultural extension can help to make smallholder farming more nutrition-sensitive. More research in this direction is needed.

Key words: agricultural extension, technology adoption, biofortification, nutrition-sensitive agriculture, Kenya

JEL codes: C93, O33, Q12, Q16, Q18

2.1 Introduction

Hunger and micronutrient malnutrition remain widespread problems in many developing countries with serious negative health consequences (FAO 2017; IFPRI 2017). Many of the people affected live in smallholder farm households. Hence, the question as to how smallholder farming can be made more nutrition-sensitive is ranking high on the development policy agenda (Pingali and Sunder 2017). The important role of market access for improving food security in the small farm sector was highlighted in recent empirical work (Bellemare and Novak 2017; Koppmair; Kassie, and Qaim 2017; Ogutu, Gödecke, and Qaim 2017; Sibhatu and Qaim 2017). In addition, agricultural technologies specifically designed to improve nutrition can possibly play an important role. Prominent examples of such pro-nutrition technologies are biofortified crops, which were bred to contain higher amounts of micronutrients, such as orange-fleshed sweet potatoes enhanced with provitamin A or high-iron rice and wheat (Bouis and Saltzman 2017; Jones and de Brauw 2015). Other examples of pro-nutrition technologies are certain species of vegetables or pulses that farmers may grow to increase household dietary diversity and address specific nutritional deficiencies (Fanzo 2017).

One problem with pro-nutrition technologies is that farmers' adoption incentives may sometimes be low (Gilligan 2012). Farmers tend to adopt new technologies rapidly when these contribute to gains in productivity and income. However, technologies that were specifically designed to improve nutrition do not necessarily increase productivity and income directly. With limited appreciation of the nutritional benefits, farmers are hesitant to adopt technologies that do not increase yield but may be associated with differences in crop taste and outward appearance. Farmers may also be concerned about not being able to market new types of crops with characteristics that are not yet widely known by traders and consumers. Even when farmers grow certain food crops primarily for home consumption, the potential to sell in the market is important when cash is needed.

Recent research showed that the adoption of pro-nutrition technologies is higher in settings where farmers have a good understanding of the technologies' agronomic and nutritional attributes (de Brauw, Eozenou, and Moursi 2015; de Brauw et al. 2015; de Groote et al. 2016). This implies that agricultural extension could and should probably play a prominent role for technology dissemination. Agricultural extension services have the mandate to facilitate technology transfer and improve innovation processes in the farming sector, but concrete experience with pro-nutrition technologies hardly exists. More generally, experience with the effectiveness of agricultural extension to promote innovation is rather mixed (Anderson and Feder 2004; Goodhue, Klonsky, and Mohapatra 2010; Läpple and Hennesy 2015). Hence, improvement in the design of agricultural extension is urgently needed. While previous studies have analyzed how agricultural training components could be improved to increase farmers' adoption of agronomic innovations (Davis et al. 2012; Läpple and Hennessy 2015), we are not aware of research that has developed and tested new extension approaches for the effective dissemination of pro-nutrition technologies. Here, we address this research gap with a randomized controlled trial (RCT) in Kenya. In particular, we evaluate how agricultural training can be combined with training in nutrition and marketing to increase farmers' adoption of a new bean variety biofortified with iron and zinc.

The name of the new bean variety is KK15. This variety was bred by the Kenya Agricultural and Livestock Research Organization (KALRO) using conventional breeding methods. Compared to other bean varieties commonly grown in Kenya, KK15 contains six times higher amounts of iron and about two times higher amounts of zinc, as a laboratory analysis that we commissioned showed. However, KK15 also differs from commonly-grown bean varieties in terms of other characteristics. According to KALRO, KK15 is high-yielding, resistant to root-rot disease, and matures earlier than most other varieties. Moreover, KK15 beans are black in color, whereas most popular bean varieties in Kenya are red. Probably because of the notable difference in outward appearance, widespread adoption of KK15 did not yet occur and may not be expected without specific extension efforts to promote this variety.

Our RCT includes three treatment arms, each with a different extension design. The first treatment only includes agricultural training. This involves explanations of the agronomic and nutritional attributes of KK15 to farmers, as well as the demonstration and training of suitable cultivation practices for this type of bean variety during different stages of the growing season. The second treatment adds specific nutritional training that goes beyond only explaining the nutritional attributes of KK15. In our study, nutrition training includes broader information about human nutritional requirements, balanced diets, and causes and consequences of nutrient deficiencies. The third treatment further adds marketing training, explaining simple mechanisms of market functioning, possible sales strategies, and linking up farmers with bean traders in the local setting. The three treatments are compared with a control group of farmers that did not receive any of these trainings, in order to evaluate the effects of the different extension designs on KK15 adoption.

The RCT was carried out in one region of Kenya and refers to one specific technology, so results cannot simply be extrapolated to other settings and technologies. Nevertheless, we expect that some broader lessons may also be learned, as evidence on the effects of combining agricultural and nutrition training is very limited. Nutrition training was shown to be an effective intervention to improve dietary quality in many situations (IFPRI 2017; Waswa et al. 2015), but such training is usually provided by nutrition and health workers, not by agricultural extension officers as in our RCT. Combining different training elements and piggybacking on existing networks of agricultural extension in rural areas could potentially be a cost-effective strategy to make smallholder farming more nutrition-sensitive.

2.2 Empirical Setting

This study builds on an RCT carried out with smallholder farmers in Western Kenya. In Kenya, smallholder agriculture accounts for nearly 75% of total agricultural production (Olwande et al. 2015). Adoption of improved agricultural technologies is relatively low among smallholders, and poverty and malnutrition are widespread (Muthayya et al. 2013; KNBS 2015; Wainana, Tongruksawattana, and Qaim 2016). The performance of extension services is mixed (Muyanga and Jayne 2008). Our RCT focuses on the adoption of a biofortified variety of beans. Kenya ranks among the top ten producers of common beans in the world (USAID 2010). In Western Kenya, most farm households cultivate beans, which are usually intercropped with maize. Beans are frequently consumed by local farm households, often on a daily basis, so that they play an important role for food security.

2.2.1 Study Region

For the study, we purposively selected two counties in Western Kenya, Kisii and Nyamira, primarily because our development partner, Africa Harvest Biotech Foundation International (Africa Harvest), had prior experience in these counties and several extension officers on the ground. Africa Harvest is a non-governmental organization and was in charge of carrying out the RCT extension treatments that we jointly designed. Given the high population density in Western Kenya, farms in Kisii and Nyamira are very small, with an average farm size of less than two acres. Farms in this region are fairly diverse and typically produce a number of food crops, such as maize, beans, sweet potatoes, bananas, and different vegetables. Many also produce cash crops such as tea and coffee and keep small herds of livestock, including chicken, sheep, goats, and sometimes cattle. Kisii and Nyamira have two agricultural seasons, the main season from March to July and a second season from September to January. However, due to favorable climatic conditions, seasonal boundaries in this part of Kenya are not very clear-cut. In terms of nutritional indicators, Kisii and Nyamira are similar to the national average. The prevalence of child stunting, the most common anthropometric measure of child undernutrition, is around 26% in both counties (KNBS 2015).

2.2.2 Sampling Strategy

Traditionally, agricultural extension was often implemented through extension officers who visited individual farmers to provide advice on specific topics (Anderson and Feder 2004). However, newer extension approaches often operate through farmer groups, which can not only increase cost-effectiveness, but also facilitate mutual learning and sharing of experiences among farmers (Davis et al. 2012; Fischer and Qaim 2012). In fact, many farmers in Kisii and Nyamira county are already organized in farmer groups registered with the Ministry of Gender, Children, and Social Development. We therefore decided to build on existing group structures and cluster the survey and the experimental treatments by farmer groups. We used a list of all existing farmer groups in Kisii and Nyamira counties, but excluded groups that had received specific development support during the previous two years to reduce possible contamination when estimating the effects of our experimental treatments. From the remaining groups on the list, we randomly selected 48 farmer groups for inclusion in the study. Of these 48 groups, 32 are located in Kisii and 16 in Nyamira county. Farmer groups in our sample have between 20 and 30 active members.

2.2.3 Farm Household Survey

In each of the 48 selected farmer groups we updated the membership lists together with the group leaders. From these membership lists, we randomly selected 20 member farmers for inclusion in the survey. However, some of the selected farmers were not available for interview, even after repeated visits. Especially in small groups it was also not always possible to replace unavailable farmers with other group members, so in some of the groups we have fewer than 20 farmers included in the survey. The survey was implemented in two rounds. The baseline round was conducted between October and December 2015, before the experimental treatments were started; it includes observations from 824 farm households. The follow-up survey was conducted between October and December 2016, after the experimental treatments were completed. Due to sample attrition, the follow-up round includes observations from 746 farm households.[1] For the evaluation, we use a balanced panel of 746 observations with complete data for both survey rounds, as this allows us to employ difference-in-difference techniques. Possible issues of attrition are addressed further below.

Data from sample households were collected through face-to-face interviews with the household head and or the spouse using a structured questionnaire. A team of agricultural students and recent graduates from the University of Nairobi assisted in carrying out the interviews in the local language after careful training. The questionnaire captured details of family demographics, agricultural production and marketing, other economic activities of the household, infrastructure and institutional conditions, and other contextual variables. Selected socioeconomic characteristics of the sample are shown in table 2.1.

Table 2.1. Selected Socioeconomic Characteristics of Sample Households at Baseline

Variables	Full sample	Treatment [a]	Control
Age of household head (years)	49.483	49.980	47.984
	(12.440)	(12.697)	(11.538)
Male household head (dummy)	0.765	0.730	0.871
	(0.424)	(0.444)	(0.336)
Education of household head (years)	8.924	8.750	9.446
	(3.732)	(3.796)	(3.490)
Farm size (acres)	1.600	1.623	1.532
	(1.253)	(1.309)	(1.067)
Number of crop and livestock species produced	12.805	12.968	12.314
	(4.625)	(4.694)	(4.387)
KK15 adopter (dummy)	0.008	0.011	0.000
	(0.089)	(0.103)	(0.000)
Observations	746	560	186

Notes: Mean values are shown with standard deviations in parentheses. [a] Treatment includes all farm households randomly assigned to one of the treatment groups.

2.3 Experimental Design

Our RCT includes three treatment groups and one control group. The 48 randomly selected farmer groups were randomly assigned to these four alternatives, 12 farmer groups each. Randomization at group level facilitates implementation of the experimental treatments and also reduces potential spillovers (Pamuk, Bulte, and Adekunle 2014).

[1] In the follow-up round of the survey we tried to reach sample households at least three times. Nevertheless, in some cases we were unable to meet respondents, because they had temporarily migrated, attended social events outside of the home community (e.g., funeral or wedding ceremonies), were sick, or had to take care of sick relatives.

2.3.1 Treatment Arms

Farmers in treatment group 1 received agricultural training, which included information about the agronomic and nutritional attributes of the KK15 bean variety and training on proper cultivation practices. Farmers in treatment group 2 received agricultural training and nutrition training. The aim of the nutrition training was to increase participants' nutrition knowledge through training on human nutritional requirements, food groups and their nutrient composition, eating balanced diets at different life stages, breast feeding practices, and health consequences of nutrient deficiencies. Farmers in treatment group 3 received agricultural training, nutrition training, and marketing training. The marketing training was aimed at enhancing participants' access to markets by increasing their knowledge on the functioning of markets and marketing strategies. It also linked farmers with bean traders through organized forums in which the characteristics of the KK15 varieties were jointly discussed. Farmers in the control group received none of these training elements during the RCT (for reasons of fairness we offered training to control group farmers in 2017, after the follow-up survey data had been collected).

2.3.2 Treatment Implementation

The trainings were administered by Africa Harvest's agricultural extension officers, who are based in the study region. In order to ensure harmonized delivery of the training contents, we did the following. First, we developed detailed manuals for each of the training components and sessions together with the extension officers. Second, we organized a workshop in which the extension officers were trained to deliver the contents with standardized methods following the manuals. This workshop also involved actual training sessions with farmer groups other than those selected for the RCT and subsequent feedback discussions in the team. Third, for the RCT we assigned extension officers to farmer groups in such a way that each officer had groups in all three treatment arms. This was important to reduce the risk of extension officer bias in evaluating the treatment effects; in spite of standardized training manuals, differences in extension officer personalities may possibly affect farmers' technology adoption behavior.

All training sessions were conducted in the regular meeting places of the farmer groups, following a structured schedule to ensure timely delivery of information. The agricultural training involved a total of seven sessions, the nutrition training involved three sessions, and the marketing training involved three sessions as well. The main training sessions were offered between January and July 2016; a summary refresher session for each of the three training components was offered in August and September 2016. Each training session lasted for about two hours.

Farmers in the treatment groups were invited to the training sessions through the group leader, who was informed and reminded of the particular date and time by the extension officers through phone calls and text messages. For all sessions, farmers and their spouses were encouraged to participate, but the decision to participate was voluntary. Participation in each of the sessions was recorded by the extension officers. In the introductory sessions, farmers were informed about the training elements and time schedule relevant for their particular treatment arm. The first sessions of all three training components (agriculture, nutrition, marketing) were conducted between January and March 2016, to be relevant for the March planting season.

Farmers who decided to adopt KK15 could place seed orders through their group leaders. Table 2.1 shows that there were a few farmers who had adopted KK15 already before the RCT started, but the adoption rate in the total sample was below 1%. As the project timeline was limited, we offered a 30% seed price subsidy to expedite the adoption process.[2] This may mean that the treatment effects are larger than they would be without the subsidy. However, as farmers in all three treatment groups and also in the control group had access to the subsidy, differences in the treatment effects on adoption can be fully attributed to the trainings and not the subsidy.

2.3.3 Covariate Balancing

Table 2.2 presents the covariate balancing tests for assessing the effectiveness of the randomization procedure in terms of delivering comparable groups. For this test we use the baseline data of households in the balanced panel. Except for very few variables where significant differences occur, the baseline characteristics are balanced across the control and treatment groups. This means that randomization bias, which is common in small samples (Barrett and Carter 2010), is not of major concern in our case. Nevertheless, to reduce any possible randomization bias, we rely on difference-in-difference estimators for evaluating the treatment effects. Moreover, we control for baseline differences in the regression models. Details of the estimation procedures are explained further below.

[2] For common varieties of beans, farmers mostly use farm-saved seeds or seeds obtained from neighbors and friends. Hence, even with the 30% subsidy, adoption of KK15 seeds was more expensive for farmers than using other varieties of beans.

Table 2.2. Mean Differences between Treatment and Control Groups at Baseline

Variables	Control – Treatment 1	Control – Treatment 2	Control – Treatment 3	Control – All Treatments
Age of household head (years)	-3.885*	-0.594	-1.437	-1.996
	(1.885)	(2.265)	(2.190)	(1.736)
Male household head (dummy)	0.113	0.193*	0.118*	0.141**
	(0.078)	(0.105)	(0.063)	(0.054)
Education of household head (years)	1.015**	0.280	0.773*	0.696**
	(0.472)	(0.559)	(0.400)	(0.332)
Household size (count)	0.473	0.379	0.536*	0.464*
	(0.348)	(0.268)	(0.279)	(0.257)
Risk attitude (scale 0 to 10)	0.136	0.062	0.510*	0.239
	(0.292)	(0.254)	(0.261)	(0.203)
Farm size (acres)	-0.088	-0.127	-0.060	-0.091
	(0.236)	(0.224)	(0.195)	(0.177)
Land title deed (dummy)	0.012	-0.044	0.017	-0.004
	(0.048)	(0.059)	(0.055)	(0.045)
Farm productive assets (1,000 Ksh)	7.962	1.738	0.241	2.061
	(9.629)	(12.655)	(13.114)	(9.730)
Own motorcycle (dummy)	-0.040	-0.003	0.012	-0.010
	(0.030)	(0.027)	(0.030)	(0.022)
Access to credit (dummy)	-0.073	0.002	0.037	-0.012
	(0.055)	(0.057)	(0.058)	(0.049)
Distance to main market (km)	-0.410	-0.841	0.633	-0.195
	(0.782)	(0.987)	(0.760)	(0.688)
Distance to extension office (km)	-0.312	-0.072	0.398	0.006
	(0.700)	(0.569)	(0.735)	(0.522)
Number of groups (count)	0.044	-0.012	0.081	0.039
	(0.073)	(0.079)	(0.067)	(0.059)
Group official (dummy)	-0.019	-0.051	0.065	-0.000
	(0.055)	(0.061)	(0.048)	(0.046)
Knows KK15 attributes (dummy) [a]	0.006	-0.013	0.011	0.002
	(0.024)	(0.029)	(0.027)	(0.021)
Knows KK15 attributes (score)	0.000	-0.013	0.004	-0.003
	(0.011)	(0.015)	(0.012)	(0.010)
KK15 adopter (dummy)	-0.005	-0.022	0.000	-0.009
	(0.005)	(0.016)	(0.000)	(0.006)
Land area under KK15 (acres)	-0.000	-0.012	0.000	-0.004
	(0.000)	(0.008)	(0.000)	(0.003)
Share of land under KK15 (%)	-0.055	-0.408	0.000	-0.150
	(0.054)	(0.297)	(0.000)	(0.102)
Seed expenditure (Ksh/acre)	424.289	-315.417	520.061	219.020
	(487.950)	(572.126)	(408.549)	(408.916)
Fertilizer expenditure (Ksh/acre)	547.114	-794.912	652.372	151.461
	(452.998)	(468.471)	(580.048)	(404.608)
Value of crop output per acre (1,000 Ksh)	1.977	-7.401	-6.865	-4.037
	(8.949)	(8.825)	(7.586)	(6.507)
Household income (1,000 Ksh)	14.548	3.321	-15.556	0.725
	(31.039)	(25.625)	(26.623)	(20.460)
Observations	376	366	376	746

Notes: Treatment 1, agricultural training. Treatment 2, agricultural training plus nutrition training. Treatment 3, agricultural training plus nutrition training plus marketing training. *, **, and *** denote significance at the 10%, 5%, and 1% level, respectively.

2.3.4 Attrition

As mentioned above, the baseline survey included 824 farm household observations, while in the follow-up survey we were only able to revisit 746 of these households. The average attrition rate is about 9%, but there is some variation across treatment and control groups (table A2.1 in the appendix). Non-random attrition might bias the randomized design and subsequently the results. Table 2.2, with data from the balanced panel, suggests that attrition did not introduce significant randomization bias. However, to be on the safe side, we test and control for attrition bias through a weighting procedure. Table A2.3 in the appendix shows probit models to analyze the association between attrition and socioeconomic variables for the baseline sample. The full-sample model in the last column of table A2.3 is used to calculate for each observation the probability to also be included in the follow-up round. These probabilities are used for inverse probability weighting in the difference-in-difference models, relying on the ignorability assumption (Wooldridge 2002).

2.3.5 Hawthorne and John Henry Effects

Apart from the treatment effects, experimental designs in randomized evaluations may potentially induce unintended behavioral changes among study participants. Changes in the behavior of the treatment group are called Hawthorne effects, while changes in the behavior of the control group are called John Henry effects (Duflo, Glennerster, and Kremer 2007). For instance, some individuals in the treatment group may be aware that they are being evaluated and may work harder to impress the evaluator. In contrast, some individuals in the control group may feel disappointed that they are not part of the treatment and either start competing with individuals in the treatment group or slack off. Such endogenous behavioral changes may lead to design contamination and possibly affect internal and external validity of the impact estimates.

We employed the following strategy to reduce possible Hawthorne and John Henry effects. First, we used cluster randomization, reducing potential behavioral change across experimental groups by limiting the likelihood of farmer groups knowing the treatments administered in other groups (Duflo et al. 2007). Second, we ensured that the household survey and the experimental treatments were implemented by different persons from different organizations to reduce the possibility of farmers drawing direct linkages between the training sessions and the household interviews. There was also no explicit mention of an evaluation during the implementation of the treatments or the survey interviews.

While farm households in the treatment groups are more likely to see the connection between the treatments and the evaluation (surveys), we feel that the risk of significant Hawthorne effects is small. The reason is that we are interested in the treatment effects on technology adoption, which is associated with a financial cost to farmers, as the KK15 seeds had to be purchased. Farmers in our sample are relatively poor. Hence, even if farmers in the treatment groups realized that they are part of an experiment, they would probably not adopt simply to impress the evaluator. A possible change in behavior might be increased attendance of the training sessions, which could possibly bias the treatment effects downward if training attendees decide not to adopt KK15 seeds. Yet we expect that even the decision to attend the training sessions will probably be made only if the expected utility from attending the training sessions is higher than the expected utility from alternative uses of time.

2.4 Estimation Strategy

We want to measure the effect of different extension treatments on farmers' adoption of the biofortified bean variety KK15. We use two indicators of technology adoption: (a) adoption of KK15 expressed as a dummy variable that takes a value of one if a household planted KK15 during the study period and zero otherwise; (b) intensity of adoption measured in terms of the percentage share of total cultivated land under KK15.

For both outcome variables, we estimate intent-to-treat (ITT) effects and treatment-on-the-treated (TOT) effects (the TOT effect is also known as local average treatment effect). The ITT effect measures the average effect of being randomly assigned into a treatment group (offer to attend certain training sessions), regardless of whether or not farmers actually attended the training sessions. The TOT effect measures the actual effect of training attendance. The ITT analysis yields precise impact estimates when there is perfect compliance, but when there is non-compliance, ITT effects get diluted and poorly predict average treatment effects (Angrist 2006). We do not observe perfect compliance in our RCT (table A2.2 in the appendix), which is why we also estimate TOT effects. The ITT effects are generally more relevant for policymakers because monitoring compliance is difficult outside experiments. On the other hand, TOT estimates are of interest to researchers to capture actual effects of the treatment itself rather than of the simple offer to be treated (Bloom 2006; Duflo et al. 2007).

For both the ITT and TOT effects, we estimate separate regression models for each of the three treatments, always with the control group observations as the reference. This allows us to compare each treatment group with the control group, while avoiding possible challenges that may arise from estimating a single regression model with multiple endogenous variables, especially in the TOT analysis.

2.4.1 Estimating Intent-To-Treat Effects

We estimate the ITT effects using the following difference-in-difference specification:

$$y_{it} = \beta_0 + \beta_1 Post_t + \beta_2 T_j + \beta_3 Post_t \times T_j + \varepsilon_{ijt}, \tag{1}$$

where y_{it} is the outcome variable of interest (KK15 adoption), $Post_t$ is a year dummy variable that takes a value of one for the follow-up data (collected in 2016), and zero for the baseline data (collected in 2015), T_j is a dummy variable that takes a value of one if the farmer group is treated, and zero otherwise (depending on the model, T_j stands for treatment group 1, treatment group 2, or treatment group 3). ε_{ijt} is the error term, clustered at farmer group level. Subscripts *i*, *t*, and *j* denote household level observation, time period, and group level observation, respectively.

The parameter of particular interest is β_3, which is the difference-in-difference estimator of the ITT effect. Under the assumption of parallel trends, which requires the difference between the control and the treatment group to remain constant over time, the difference-in-difference estimator overcomes possible selection bias from the absence of perfect balance in the baseline covariates. This estimator also accounts for time-invariant unobserved heterogeneity (Greene 2012). Equation (1) is estimated with ordinary least squares (OLS). For the binary adoption outcome we use a linear probability model (LPM). While the LPM may generate predicted probabilities outside the unit interval, its marginal effects are generally close to those from non-linear models (Angrist and Pischke 2009).

To control for differences in baseline covariates, we extend the model in equation (1) as follows:

$$y_{it} = \beta_0 + \beta_1 Post_t + \beta_2 T_j + \beta_3 Post_t \times T_j + \delta \mathbf{x_{ij}} + \varepsilon_{ijt}, \tag{2}$$

where $\mathbf{x_{ij}}$ is a vector of socioeconomic controls.

2.4.2 Estimating Treatment-On-The-Treated Effects

To estimate the TOT effects, we use actual training attendance as the treatment variable. Since several training sessions were offered and it is possible that farmers participated in some but not all of these sessions, we measure training attendance in two different ways: (a) a dummy variable that takes a value of one if a household attended at least one of the training sessions that were offered in his/her group, and zero otherwise; (b) intensity of training attendance, measured by the number of training sessions attended relative to all training sessions offered in the group (this share can take values between zero and one).

The decision to attend training sessions is endogenous. To avoid endogeneity bias we use an instrumental variables (IV) approach, relying on the random assignment into the treatment groups (offer to attend certain trainings) as a valid instrument for training attendance. Using the randomization status as an instrument is a common approach in the RCT literature (Ashraf, Giné, and Karlan 2009; Carter, Laajaj, and Yang 2013). The TOT effect estimates are unbiased under the following assumptions (Angrist, Imbens, and Rubin 1996; Angrist and Pischke 2009; Ashraf et al. 2009): First, the offer to participate in the treatment is random, which is fulfilled in our case due to random assignment of farmer groups to different treatments. Second, the offer to participate in the treatment is highly correlated with actual training attendance, which is also fulfilled in our case. Third, the offer to participate in the treatment is not correlated with the outcome variables, except through actual attendance of the training sessions. This third assumption is more challenging to test; it can be violated if there are within-group externalities, for instance, if the behavior of non-attendees in the training sessions is affected by the behavior of attendees. Farmer groups are usually designed to facilitate cooperation among members, so that within-group externalities may occur. We will therefore interpret the TOT effect estimates cautiously. However, it is important to note that within-group externalities – if existent – would lead to a downward bias, meaning that the true TOT effects could be larger than the ones estimated with the IV approach.

We estimate the TOT effects using the following IV difference-in-difference specification:

$$y_{it} = \alpha_0 + \alpha_1 Post_t + \alpha_2 \hat{T}_i + \alpha_3 Post_t \times \hat{T}_i + \upsilon_{ijt}, \qquad (3)$$

where $\hat{T}_i$ is the fitted value of the treatment (actual training attendance) obtained from the first-stage regression with the instrument. α_3 is the parameter of interest, and υ_{ijt} is the error term, clustered at farmer group level.

Again, to control for differences in baseline covariates, we extend the model in equation (3) as follows:

$$y_{it} = \alpha_0 + \alpha_1 Post_t + \alpha_2 \hat{T}_i + \alpha_3 Post_t \times \hat{T}_i + \delta \mathbf{x}_{ij} + \upsilon_{ijt}. \qquad (4)$$

For the estimation of the models in equations (3) and (4) we apply two-stage least squares (2SLS). Non-linear models, such as IV probit and Tobit could have been used, but these require the endogenous regressors to be continuous (StataCorp 2013). The 2SLS estimator works efficiently and produces estimates with a robust causal interpretation also with limited dependent variables (Angrist 2006).

2.5 Estimation Results

We now present and discuss the results of our analysis following the estimation strategy explained in the previous section.

2.5.1 Intent-To-Treat Effects on Technology Adoption

In table 2.3, we present estimates of the ITT effects for the decision to adopt KK15 bean seeds, as well as for adoption intensity (share of land under KK15). We show models with and without attrition-weighting. The results between both alternatives are similar. In the discussion, we focus on the attrition-weighted results. For each model, we also show estimates with and without baseline controls included: the ITT effects in both specifications are identical, suggesting that the difference-in-difference procedure controls for baseline differences very well.

The results in table 2.3 show positive and significant effects of all three treatments on the likelihood of KK15 adoption, and also on adoption intensity, suggesting that the extension approaches are effective in terms of increasing the uptake of this pro-nutrition technology. The attrition-weighted ITT estimates in panel (A) of table 2.3 imply that farmers who were offered agricultural training alone (treatment 1) are 22.5 percentage points more likely to plant KK15 seeds than their colleagues in the control group. The share of land under KK15 is 4.9 percentage points higher. For farmers who were offered agricultural training and nutrition training (treatment 2 shown in panel B of table 2.3), the likelihood of planting KK15 seeds is 26 percentage points higher than for farmers in the control group. That is, the nutrition training seems to further increase technology adoption over and above the effect of agricultural training alone. However, farmers in treatment 3 (panel C) have a slightly lower likelihood of KK15 adoption than farmers in treatments 1 and 2.

Table 2.3. Effects of Extension Treatments on Technology Adoption, Intent-To-Treat Estimates

Variables	Results without attrition-weighting				Attrition-weighted results			
	Planted KK15 (dummy)		Share of land under KK15 (%)		Planted KK15 (dummy)		Share of land under KK15 (%)	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: Treatment 1 (n=752)								
Post (dummy)	0.005	0.005	0.045	0.045	0.004	0.004	0.037	0.037
	(0.005)	(0.005)	(0.045)	(0.045)	(0.004)	(0.004)	(0.037)	(0.037)
Treatment 1 (dummy)	0.005	0.008	0.055	0.177	0.006	0.011	0.058	0.142
	(0.005)	(0.011)	(0.054)	(0.208)	(0.005)	(0.017)	(0.053)	(0.437)
Post × Treatment 1	0.237***	0.237***	4.832***	4.832***	0.225**	0.225**	4.929**	4.929**
	(0.074)	(0.075)	(1.575)	(1.587)	(0.074)	(0.082)	(1.989)	(2.004)
Baseline controls	No	Yes	No	Yes	No	Yes	No	Yes
R-squared	0.184	0.199	0.113	0.124	0.163	0.175	0.096	0.113
Panel B: Treatment 2 (n=732)								
Post (dummy)	0.005	0.005	0.045	0.045	0.004	0.004	0.037	0.037
	(0.005)	(0.005)	(0.045)	(0.045)	(0.004)	(0.004)	(0.037)	(0.037)
Treatment 2 (dummy)	0.022	0.016	0.409	0.274	0.023	0.020	0.429	0.393
	(0.017)	(0.012)	(0.298)	(0.230)	(0.017)	(0.013)	(0.300)	(0.245)
Post × Treatment 2	0.234***	0.234***	4.420***	4.420***	0.261***	0.261***	4.814***	4.814***
	(0.060)	(0.061)	(1.173)	(1.182)	(0.075)	(0.075)	(1.318)	(1.328)
Baseline controls	No	Yes	No	Yes	No	Yes	No	Yes
R-squared	0.180	0.198	0.106	0.118	0.207	0.227	0.118	0.129
Panel C: Treatment 3 (n=752)								
Post (dummy)	0.005	0.005	0.045	0.045	0.004	0.004	0.037	0.037
	(0.005)	(0.005)	(0.045)	(0.045)	(0.004)	(0.004)	(0.037)	(0.037)
Treatment 3 (dummy)	0.022	0.002	0.044	-0.052	0.005	0.004	0.042	-0.002
	(0.017)	(0.011)	(0.043)	(0.195)	(0.005)	(0.011)	(0.041)	(0.212)
Post × Treatment 3	0.200***	0.200***	3.871***	3.871***	0.214***	0.214***	4.443***	4.443***
	(0.047)	(0.048)	(1.123)	(1.124)	(0.052)	(0.052)	(1.454)	(1.465)
Baseline controls	No	Yes	No	Yes	No	Yes	No	Yes
R-squared	0.154	0.177	0.093	0.114	0.165	0.192	0.104	0.125

Notes: Coefficient estimates are shown with robust standard errors clustered at farmer group level in parentheses. Post, dummy variable which takes a value of 1 for follow-up round observations (after treatment), and zero for baseline observations. Treatment 1, agricultural training. Treatment 2, agricultural training plus nutrition training. Treatment 3, agricultural training plus nutrition training plus marketing training. Baseline controls include age, gender, education, risk attitude, household size, farm size, value of productive assets, access to credit, distance to market, group official, and county dummy.*, **, and *** denote significance at the 10%, 5%, and 1% level, respectively.

As explained, the ITT results measure the effects of the training offers, without looking at farmers' actual attendance of training sessions. The effects of actual attendance are analyzed in the following.

2.5.2 Treatment-On-The-Treated Effects on Technology Adoption

Tables 2.4 and 2.5 present the estimated TOT effects. These estimates are computed by comparing farmers who attended the training sessions offered with those who did not attend, including the control group. Table 2.4 shows results of models where the attendance dummy is used as the treatment variable. The attrition-weighted results suggest that farmers who attended agricultural training only (treatment 1 shown in panel A of table 2.4) are 22.5 percentage points more likely to adopt KK15 beans than their colleagues who did not attend any of the training sessions. This refers to the model with baseline controls (column 6). The adoption intensity is 4.9 percentage points higher (column 8).

Farmers who attended agricultural and nutrition training (treatment 2 shown in panel B of table 2.4) are 32.2 percentage points more likely to adopt KK15 than those who did not attend any of the trainings. Their adoption intensity is 6.1 percentage points higher. The comparison of the TOT effects between treatment 1 and treatment 2 suggests that attendance of nutrition training increases KK15 adoption by 9.7 percentage points and adoption intensity by 1.2 percentage points over and above attendance of agricultural training alone. The TOT effects in panel (C) of table 2.4 are very similar to those in panel (B), which may imply that attending marketing training may not have an additional effect on adoption over and above agricultural and nutrition training.

Table 2.5 shows results of the TOT effects models with intensity of training attendance as the treatment variable. As explained, intensity is measured in terms of the share of training sessions attended with values ranging between zero and one. The attrition-weighted estimates with baseline controls in panel (A) suggest that farmers who attended all of the agricultural training sessions offered in treatment 1 are 40.9 percentage points more likely to adopt KK15 beans than farmers who did not attend any of the training sessions. Full attendance of all agricultural training sessions increases the adoption intensity by 7.3 percentage points.

In panel (B) of table 2.5, we observe that farmers who attended all of the agricultural and nutrition training sessions offered in treatment 2 are 52.8 percentage points more likely to adopt KK15 beans than their colleagues who attended none of the sessions.

Table 2.4. Effects of Extension Treatments on Technology Adoption, Treatment-On-The-Treated Estimates (IV Results with Training Attendance Dummies as Treatment Variables)

	Results without attrition-weighting				Attrition-weighted results			
Variables	Planted KK15 (dummy)		Share of land under KK15 (%)		Planted KK15 (dummy)		Share of land under KK15 (%)	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: Treatment 1 (n=752)								
Post (dummy)	0.022	0.024*	0.387	0.475	0.025	0.027	0.454	0.577
	(0.013)	(0.014)	(0.275)	(0.331)	(0.016)	(0.017)	(0.323)	(0.403)
Treatment 1 (dummy)	0.026*	0.032	0.477	0.740	0.026*	0.035	0.464	0.698
	(0.015)	(0.020)	(0.296)	(0.512)	(0.015)	(0.029)	(0.288)	(0.750)
Post × Treatment 1	0.247***	0.242***	5.028***	4.817***	0.229***	0.225**	5.093**	4.861**
	(0.073)	(0.074)	(1.550)	(1.552)	(0.085)	(0.088)	(2.122)	(2.155)
Baseline controls	No	Yes	No	Yes	No	Yes	No	Yes
R-squared	0.212	0.224	0.130	0.138	0.190	0.200	0.113	0.127
Panel B: Treatment 2 (n=732)								
Post (dummy)	0.020	0.016	0.231	0.144	0.015*	0.013	0.173	0.127
	(0.012)	(0.010)	(0.148)	(0.131)	(0.009)	(0.008)	(0.109)	(0.099)
Treatment 2 (dummy)	0.051	0.040	0.822	0.572	0.047	0.043	0.771*	0.738*
	(0.034)	(0.026)	(0.512)	(0.459)	(0.030)	(0.024)	(0.464)	(0.437)
Post × Treatment 2	0.280***	0.291***	5.553***	5.797***	0.316***	0.322***	6.014***	6.149***
	(0.070)	(0.073)	(1.571)	(1.720)	(0.084)	(0.088)	(1.669)	(1.792)
Baseline controls	No	Yes	No	Yes	No	Yes	No	Yes
R-squared	0.246	0.266	0.151	0.164	0.279	0.299	0.163	0.175
Panel C: Treatment 3 (n=752)								
Post (dummy)	0.003	0.002	0.022	-0.006	0.002	0.002	0.019	0.008
	(0.003)	(0.004)	(0.023)	(0.089)	(0.002)	(0.004)	(0.019)	(0.097)
Treatment 3 (dummy)	0.004	-0.001	0.032	-0.120	0.004	0.003	0.035	-0.033
	(0.009)	(0.021)	(0.071)	(0.392)	(0.008)	(0.020)	(0.064)	(0.421)
Post × Treatment 3	0.307***	0.310***	5.858***	5.943***	0.317***	0.317***	6.500***	6.533***
	(0.063)	(0.067)	(1.635)	(1.720)	(0.066)	(0.069)	(1.842)	(1.905)
Baseline controls	No	Yes	No	Yes	No	Yes	No	Yes
R-squared	0.259	0.281	0.156	0.175	0.268	0.293	0.167	0.187

Notes: Coefficient estimates are shown with robust standard errors clustered at farmer group level in parentheses. Post, dummy variable which takes a value of 1 for follow-up round observations (after treatment), and zero for baseline observations. Treatment 1, agricultural training. Treatment 2, agricultural training plus nutrition training. Treatment 3, agricultural training plus nutrition training plus marketing training. Baseline controls include age, gender, education, risk attitude, household size, farm size, value of productive assets, access to credit, distance to market, group official, and county dummy. *, **, and *** denote significance at the 10%, 5%, and 1% level, respectively.

Table 2.5. Effects of Extension Treatments on Technology Adoption, Treatment-On-The-Treated Estimates (IV Results with Intensity of Training Attendance as Treatment Variables)

	Results without attrition-weighting				Attrition-weighted results			
Variables	Planted KK15 (dummy)		Share of land under KK15 (%)		Planted KK15 (dummy)		Share of land under KK15 (%)	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: Treatment 1 (n=752)								
Post (dummy)	0.021	0.023	0.788	0.853*	0.009	0.012	0.678	0.749
	(0.021)	(0.022)	(0.487)	(0.513)	(0.026)	(0.027)	(0.596)	(0.642)
Treatment 1 (share)	0.040	0.052	1.552*	1.959*	0.018	0.036	1.089	1.447
	(0.040)	(0.050)	(0.794)	(1.065)	(0.041)	(0.062)	(1.842)	(1.450)
Post × Treatment 1	0.405***	0.396***	6.599***	6.346***	0.417***	0.409***	7.527***	7.309**
	(0.117)	(0.120)	(2.129)	(2.142)	(0.136)	(0.148)	(2.792)	(3.017)
Baseline controls	No	Yes	No	Yes	No	Yes	No	Yes
R-squared	0.270	0.278	0.141	0.147	0.261	0.272	0.135	0.150
Panel B: Treatment 2 (n=732)								
Post (dummy)	0.025	0.021	0.989	0.893	0.023	0.021	1.043	0.988
	(0.025)	(0.023)	(0.755)	(0.698)	(0.024)	(0.022)	(0.781)	(0.721)
Treatment 2 (share)	0.107	0.969	3.460	3.188	0.103	0.108	3.615*	3.734**
	(0.080)	(0.739)	(2.096)	(1.942)	(0.072)	(0.066)	(2.048)	(1.868)
Post × Treatment 2	0.488***	0.510***	6.318***	6.812***	0.518***	0.528***	6.109***	6.392***
	(0.102)	(0.117)	(2.223)	(2.401)	(0.113)	(0.130)	(2.091)	(2.219)
Baseline controls	No	Yes	No	Yes	No	Yes	No	Yes
R-squared	0.319	0.341	0.144	0.160	0.348	0.370	0.145	0.159
Panel C: Treatment 3 (n=752)								
Post (dummy)	0.019	0.020	0.495	0.512	0.016	0.018	0.638	0.660
	(0.020)	(0.018)	(0.442)	(0.476)	(0.020)	(0.019)	(0.563)	(0.594)
Treatment 3 (share)	0.056	0.058	1.474	1.407	0.048	0.053	1.823	1.828
	(0.056)	(0.054)	(1.278)	(1.643)	(0.052)	(0.055)	(1.544)	(1.806)
Post × Treatment 3	0.521***	0.514***	8.988***	8.884**	0.550***	0.540***	9.422**	9.294**
	(0.116)	(0.119)	(3.464)	(3.783)	(0.124)	(0.131)	(3.809)	(4.127)
Baseline controls	No	Yes	No	Yes	No	Yes	No	Yes
R-squared	0.330	0.344	0.182	0.196	0.354	0.370	0.190	0.205

Notes: Coefficient estimates are shown with robust standard errors clustered at farmer group level in parentheses. Post, dummy variable which takes a value of 1 for follow-up round observations (after treatment), and zero for baseline observations. Treatment 1, agricultural training. Treatment 2, agricultural training plus nutrition training. Treatment 3, agricultural training plus nutrition training plus marketing training. Baseline controls include age, gender, education, risk attitude, household size, farm size, value of productive assets, access to credit, distance to market, group official, and county dummy. *, **, and *** denote significance at the 10%, 5%, and 1% level, respectively.

This TOT effect on adoption is almost 12 percentage points higher than that in treatment 1, further supporting the hypothesis that nutrition training components can increase the effectiveness of extension for pro-nutrition technologies. In terms of the adoption intensity, this difference is not visible.

In panel (C) of table 2.5, the treatment effect on adoption is similar as in panel (B), suggesting that attending the additional marketing training does not make a major difference for the adoption decision. However, the TOT effect on adoption intensity in treatment 3 is somewhat higher than in the other two treatments. This makes sense, because marketing training is particularly relevant when a marketable surplus is produced, which is more likely when a larger share of the farm area is cultivated with KK15 beans.

As expected, most of the estimated TOT effects are larger than the ITT effects, which is especially true when using the intensity of training attendance as the treatment variable. This comparison implies that the ITT estimates were affected by non-compliance problems.

2.5.3 Heterogeneous Treatment Effects

We examine heterogeneous treatment effects by gender and education of the farmer (household head). Gender and education were shown to be important variables in analyses of technology adoption and the effectiveness of agricultural extension in the African small farm sector (Anderson and Feder 2004; Kabunga, Dubois, and Qaim 2012; Lambrecht et al. 2014). For education, we create an education dummy variable that takes a value of one if the farmer had at least eight years of education (i.e., post-primary education), and zero otherwise. Our sample size does not allow us to carry out the analysis with more than these two education categories. Table A2.4 in the appendix shows ITT models with additional interaction terms between the treatment and the education dummy. For treatments 1 and 2, all interaction terms are statistically insignificant, suggesting that the treatment effects do not differ by education status.

Many previous studies showed that farmers with higher levels of education are better able to absorb new information and are more likely to be early adopters of new technologies (Fisher and Kandiwa 2014; Foster and Rosenzweig 2010; Wainaina et al. 2016). Hence, our result of homogenous treatment effects by education may surprise. However, it should be noted that KK15 is a technology that is not very difficult to understand and implement, as most farmers were already familiar with growing beans. Also, the training sessions offered in our RCT were tailored to farmers with relatively low levels of education: the extension officers used local dialects to explain concepts, employed visual aids such as posters and flipcharts, provided practical demonstrations, and moderated interactive question and answer sessions. These methods facilitated understanding also for farmers with low levels of education. The only treatment where positive interactions between the treatment and the education dummy are observed in table A2.4 is treatment 3, suggesting that the marketing training sessions may have been a bit more difficult to understand for farmers with low levels of education.

Table A2.5 in the appendix analyzes heterogeneous treatment effects by gender. None of the interactions between the treatment and the gender dummy are statistically significant, which suggests that the treatment effects do not differ between male and female farmers. Previous studies showed that women farmers are often slower or less likely to adopt new agricultural technologies due to various constraints (Fisher and Kandiwa 2014; Peterman, Behrman, and Quisumbing 2014). However, a major constraint for women in technology adoption is access to proper information and extension (Kabunga et al. 2012). This was not an issue in our RCT. Around one-quarter of the farmers in our sample were female, and these female farmers were as likely as their male colleagues to attend the training sessions offered in the treatment groups.[3]

2.5.4 How Does Training Influence Adoption?

The experimental results show that the training sessions offered have influenced farmers' decision to adopt KK15 in a positive way. We hypothesize that the treatment effects of training on adoption are channeled through farmers' increased awareness of KK15 and knowledge about this variety's attributes and their wider implications. We test this hypothesis using a simple knowledge score computed based on the proportion of farmers' correct responses to questions on the technology's attributes. The same questions were asked during the baseline and follow-up surveys. Farmers were first asked about their awareness of the KK15 bean variety. Only farmers who reported that they were aware of KK15 were subsequently asked several questions to assess their knowledge of the technology's attributes. These questions related to the agronomic characteristics of KK15 as well to the variety's nutritional value. Farmers who were unaware of the technology were automatically assigned a zero value in the knowledge score.

[3] We also estimated heterogeneous treatment effects by education and gender using the TOT models. The interaction terms were not statistically significant in any of these models.

Results of different regression models with this knowledge score are shown in table 2.6. In column (1), we show results of a simple probit model of KK15 adoption, where the knowledge score is included as an explanatory variable. As expected, higher levels of knowledge about KK15 contribute to higher levels of adoption.

Table 2.6. Knowledge about KK15 Attributes and Technology Adoption

Variables	Probit	IV Probit	
	Planted KK15	KK15 Score First-stage	Planted KK15 Second-stage
	(1)	(2)	(3)
KK15 knowledge score (0-1)	0.116***		0.836***
	(0.026)		(0.250)
Post (dummy)	0.025**	0.374***	-0.173*
	(0.011)	(0.036)	(0.091)
Age of household head (years)	0.001***	-0.000	0.002**
	(0.000)	(0.001)	(0.001)
Male household head (dummy)	-0.001	-0.018	0.016
	(0.010)	(0.021)	(0.017)
Education of household head (years)	-0.001	0.001	-0.002
	(0.001)	(0.002)	(0.003)
Household size (number)	0.001	0.001	0.001
	(0.001)	(0.003)	(0.004)
Farm size (acres)	0.002	0.007	0.000
	(0.002)	(0.006)	(0.006)
Farm productive assets (1,000 Ksh)	-0.000	-0.000	0.000
	(0.000)	(0.000)	(0.000)
Access to credit (dummy)	0.001	0.028	-0.016
	(0.010)	(0.018)	(0.022)
Distance to closest market (km)	0.000	0.001	0.001
	(0.001)	(0.001)	(0.001)
Group official (dummy)	0.005	0.052***	-0.020
	(0.007)	(0.016)	(0.022)
Kisii County (dummy) [a]	0.011	-0.026	0.038*
	(0.011)	(0.024)	(0.022)
Treatment assignment (dummy) [b]		0.219***	
		(0.016)	
Constant		-0.175***	
		(0.053)	
Pseudo R-squared	0.431		
Log pseudo-likelihood	-265.777		-552.547
Wald χ^2			1231.230***
Wald test of exogeneity			16.530***
Observations	1492	1492	1492

Notes: Coefficient estimates are shown with standard errors in parentheses. Estimates in columns (1) and (3) are marginal effects. [a] Nyamira County is the reference. [b] Assignment to any of the three treatment groups. *, **, and *** significant at 10%, 5%, and 1% level, respectively.

However, knowledge about KK15 is potentially endogenous, so we also estimate an IV probit model, using the random treatment assignment as an instrument for the KK15 knowledge score. Column (2) of table 2.6 shows the first-stage results of the IV probit, which confirm that the training sessions offered in the experimental treatments contributed to increasing farmers' knowledge scores. The second-stage results of the IV probit are presented in column (3). They show that knowledge about the attributes of KK15 increases adoption significantly also after controlling for endogeneity. The estimated marginal effect suggests that full knowledge about KK15 would increase the probability of adoption by 84 percentage points. This large effect clearly underlines the importance of effective information and knowledge flows for technology adoption among smallholder farmers.

2.6 Conclusion

In this article, we have analyzed how agricultural extension can be improved to increase the adoption of pro-nutrition technologies by smallholder farm households. In particular, we have studied how agricultural training can be combined with nutrition training and marketing training to increase the adoption of KK15, a new variety of beans biofortified with iron and zinc. Different extension treatments were implemented in an RCT with smallholder farm households in Western Kenya. Treatment effects were estimated with difference-in-difference models, using data from baseline and follow-up surveys.

Results show that intensive training offered by agricultural extension officers and tailored to local conditions can increase technology adoption considerably within a relatively short period of time. In all three treatments, the adoption of KK15 increased from less than 1% before the RCT started to more than 20% one year later. This rapid increase in adoption in the treatment groups suggests that farmers are willing to adopt pro-nutrition technologies, when they are well informed about the attributes and their implications, even when the technologies are not primarily designed to increase productivity and income. Even though farmers in the RCT received a 30% subsidy on the price of KK15 seeds, they had to pay for the seeds and therefore made a real adoption decision considering expected benefits and costs.

Comparison of the different treatments revealed interesting additional insights. Farmers who had received agricultural training and nutrition training were more likely to adopt KK15 than farmers who had only received agricultural training. Comparison of the TOT effects suggests that additional nutrition training further increased adoption rates by 10-12 percentage points over and above the effects of agricultural training alone. This additional effect of nutrition training may not surprise, because of the positive nutritional attributes of KK15. However, it should be noted that these attributes of KK15 were communicated to farmers in the agricultural training sessions. The nutrition training sessions covered broader aspects related to healthy nutrition, balanced diets, and the health consequences of nutrient deficiencies. It seems that knowledge about these broader nutrition aspects has helped farmers to better appreciate the nutrition attributes of KK15, thus resulting in higher adoption rates. The nutrition training may certainly have positive effects on household diets and health beyond KK15 adoption. Analysis of such wider effects is beyond the scope of this study.

Our findings have important policy implications. Nutrition education is usually not delivered through the agricultural extension service, but through specialized nutrition and health workers. Our results suggest that combining agricultural and nutrition training in agricultural extension approaches is feasible. Of course, the nutrition training should be designed together with nutrition experts, and the agricultural extension officers first need to be trained before they can effectively deliver nutrition training to farm families. However, the high personnel and logistics cost of reaching out to families in rural areas is a major impediment for more widespread coverage of nutrition and health education campaigns. Based on our results we argue that closer cooperation between agricultural extension and nutrition and health organizations can be a cost-effective way to promote pro-nutrition innovations among smallholder farm households.

The additional marketing training provided in one of the treatment arms of our RCT did not contribute to higher KK15 adoption over and above the effects of agricultural and nutrition training. This is surprising because research has shown that improved market access can improve technology adoption in the small farm sector (Fischer and Qaim 2012b). That the marketing training did not have an additional effect in our study may be due to the fact that we only considered adoption during the one year in which the training sessions were implemented. During this very early period of KK15 adoption, most of the adopting farmers planted small areas with the new variety, in order to test out the technology's attributes. The small quantities harvested were primarily consumed at home and not marketed. It is possible that the marketing training will have larger effects when farmers consider increasing the area cultivated with KK15 at a later stage. Indeed, in some of the TOT models we found that the marketing treatment had a significantly positive additional effect on adoption intensity.

The study region in Western Kenya with very small farm sizes, diverse production systems, limited market access due to infrastructure constraints, and relatively high rates of malnutrition is typical for the African small farm sector. Hence, some of the general findings will also be relevant beyond this specific setting. However, the exact estimates of the treatment effects should not be generalized. There are particularly two factors in our RCT that may reduce the external validity of the empirical estimates. First, our extension treatments were fairly intense. Within a period of nine months, farmers in all treatment groups were offered seven agricultural training sessions. In some of the treatment groups, three nutrition training and three marketing training sessions were offered in addition. Outside an experiment, the training frequency and intensity may be lower, meaning that the effects on technology adoption may be lower too. Second, we only analyzed the short-term adoption effects, as the follow-up survey was carried out less than one year after the treatments had started. Technology adoption is a process over time. Most farmers seemed to be satisfied with KK15 during the first year of adoption, so it is likely that adoption rates will further increase in the future, among both treated and untreated farmers. Further research is needed on how the design of agricultural extension approaches can be improved in order to increase the adoption of pro-nutrition technologies. Our study is only an initial step in this direction.

Appendix A2

Table A2.1. Attrition Rates Across Treatment and Control Groups

	Baseline number (%)	Follow-up number (%)	Attriting number (%)
Control	209 (25.36)	186 (24.93)	23 (11.00)
Treatment 1	204 (24.76)	190 (25.47)	14 (6.86)
Treatment 2	208 (25.24)	180 (24.13)	28 (13.46)
Treatment 3	203 (24.64)	190 (25.47)	13 (6.40)
Observations	824 (100)	746 (100)	78 (9.47)

Notes: Treatment 1, agricultural training. Treatment 2, agricultural training plus nutrition training. Treatment 3, agricultural training plus nutrition training plus market training.

Table A2.2. Compliance Rates with Training Attendance

	Mean	SD	Observations
Households attending treatment 1 (dummy)	0.826	0.380	190
Households attending treatment 2 (dummy)	0.728	0.446	180
Households attending treatment 3 (dummy)	0.668	0.472	190
Share of total trainings in treatment 1 attended	0.509	0.353	190
Share of total trainings in treatment 2 attended	0.396	0.348	180
Share of total trainings in treatment 3 attended	0.332	0.345	190

Notes: Treatment 1, agricultural training. Treatment 2, agricultural training plus nutrition training. Treatment 3, agricultural training plus nutrition training plus market training.

Table A2.3. Attrition Probit Regressions

Variables	Treatment 1 Attrition	Treatment 2 Attrition	Treatment 3 Attrition	Control Attrition	Full sample Attrition
Treatment 1					-0.188
					(0.182)
Treatment 2					0.140
					(0.149)
Treatment 3					-0.299**
					(0.135)
Age of household head (years)	-0.219*	-0.048	-0.075*	-0.106*	-0.086**
	(0.118)	(0.118)	(0.044)	(0.055)	(0.036)
Age-squared (years)	0.002	0.000	0.001	0.001*	0.001**
	(0.001)	(0.001)	(0.000)	(0.001)	(0.000)
Male household head (dummy)	-0.496	0.340	0.230	0.490	0.188
	(0.424)	(0.410)	(0.223)	(0.384)	(0.173)
Education of household head (years)	-0.024	-0.012	0.014	-0.007	-0.002
	(0.045)	(0.041)	(0.050)	(0.053)	(0.023)
Household size (count)	0.026	0.065	-0.159	0.078	0.023
	(0.108)	(0.054)	(0.106)	(0.063)	(0.034)
Risk attitude (scale 0 to 10)	-0.086	0.027	0.068	-0.068	0.003
	(0.056)	(0.049)	(0.060)	(0.053)	(0.023)
Farm size (acres)	0.228**	0.154*	-0.062	-0.030	0.053
	(0.101)	(0.090)	(0.104)	(0.153)	(0.053)
Land title deed (dummy)	0.099	-0.014	0.314	0.510*	0.175
	(0.388)	(0.177)	(0.358)	(0.275)	(0.124)
Farm productive assets (1,000 Ksh)	-0.007	0.001	0.001	-0.003*	0.001
	(0.008)	(0.001)	(0.001)	(0.002)	(0.000)
Own motorcycle (dummy)	0.039	-0.629	-0.768	-0.329	-0.438**
	(0.694)	(0.550)	(0.559)	(0.407)	(0.214)
Access to credit (dummy)	-0.034	-0.087	-0.342	0.291	-0.024
	(0.476)	(0.267)	(0.383)	(0.350)	(0.151)
Distance to main market (km)	0.034**	0.023	0.010	-0.003	0.020***
	(0.017)	(0.014)	(0.046)	(0.018)	(0.008)
Distance to extension office (km)	-0.018	-0.032	-0.069**	-0.001	-0.031**
	(0.022)	(0.025)	(0.028)	(0.026)	(0.015)
Number of groups (count)	0.539***	-0.045	0.465*	-0.710**	0.083
	(0.167)	(0.156)	(0.265)	(0.316)	(0.106)
Group official (dummy)	-0.145	-0.133	0.242	-0.884**	-0.198
	(0.351)	(0.287)	(0.286)	(0.410)	(0.161)
Off-farm income (dummy)	0.565	0.249	0.047	-0.249	0.092
	(0.546)	(0.324)	(0.337)	(0.407)	(0.177)
Seed expenditure per acre	0.000	-0.000	0.000	-0.000**	-0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Fertilizer expenditure per acre	0.000***	0.000	-0.000	-0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Kisii County [a]	1.086***	-0.411	-0.019	0.099	-0.046
	(0.383)	(0.312)	(0.295)	(0.330)	(0.141)
Constant	3.080	-0.683	0.216	2.468	0.561
	(2.288)	(2.788)	(1.572)	(1.409)	(0.830)
Pseudo R-squared	0.320	0.108	0.177	0.213	0.074
Observations	204	208	203	209	824

Notes: Coefficient estimates are shown with robust standard errors clustered at farmer group level in parentheses. Treatment 1, agricultural training. Treatment 2, agricultural training plus nutrition training. Treatment 3, agricultural training plus nutrition training plus market training. [a] Nyamira County is the reference. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

Table A2.4. Effects of Extension Treatments on Technology Adoption, Intent-To-Treat Estimates by Education

Variables	Results without attrition-weighting				Attrition-weighted results			
	Planted KK15 (dummy)[a]		Share of land under KK15 (%)		Planted KK15 (dummy)		Share of land under KK15 (%)	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: Treatment 1 (n=752)								
Post (dummy)	0.012	0.012	0.098	0.098	0.010	0.010	0.081	0.081
	(0.012)	(0.012)	(0.097)	(0.099)	(0.010)	(0.010)	(0.082)	(0.082)
Education (dummy)	0.000	-0.007	0.000***	-0.245	-0.000	-0.013	0.000	-0.495
	(0.000)	(0.006)	(0.000)	(0.204)	(0.000)	(0.009)	(.)	(0.402)
Treatment 1 (dummy)	0.010	0.013	0.103	0.263	0.010	0.019	0.103	0.270
	(0.009)	(0.017)	(0.098)	(0.314)	(0.009)	(0.022)	(0.091)	(0.570)
Post × Education	-0.012	-0.012	-0.098	-0.098	-0.010	-0.010	-0.081	-0.081
	(0.012)	(0.012)	(0.098)	(0.099)	(0.010)	(0.010)	(0.082)	(0.082)
Post × Treatment 1	0.233**	0.233**	4.868**	4.868**	0.223*	0.223*	5.071**	5.071*
	(0.096)	(0.097)	(1.850)	(1.863)	(0.110)	(0.112)	(2.438)	(2.472)
Post × Treatment 1 × Education	0.005	0.005	-0.095	-0.095	0.002	0.002	-0.343	-0.343
	(0.079)	(0.079)	(1.221)	(1.230)	(0.091)	(0.092)	(1.524)	(1.535)
Baseline controls	No	Yes	No	Yes	No	Yes	No	Yes
R-squared	0.185	0.200	0.113	0.125	0.163	0.177	0.096	0.115
Panel B: Treatment 2 (n=732)								
Post (dummy)	0.012	0.012	0.098	0.098	0.010	0.010	0.081	0.081
	(0.012)	(0.012)	(0.098)	(0.099)	(0.010)	(0.010)	(0.082)	(0.082)
Education (dummy)	0.000	-0.004	0.000	-0.121	0.000	-0.006	0.000	-0.170
	(.)	(0.008)	(.)	(0.153)	(.)	(0.009)	(0.000)	(0.181)
Treatment 2 (dummy)	0.048	0.042	0.886	0.767	0.049	0.043	0.898	0.832
	(0.035)	(0.027)	(0.639)	(0.456)	(0.034)	(0.028)	(0.612)	(0.412)
Post × Education	-0.012	-0.012	-0.098	-0.098	-0.010	-0.010	-0.081	-0.081
	(0.012)	(0.012)	(0.098)	(0.099)	(0.010)	(0.010)	(0.082)	(0.082)
Post × Treatment 2	0.229**	0.229**	3.953**	3.953**	0.267**	0.267**	4.395**	4.395**
	(0.088)	(0.088)	(1.588)	(1.599)	(0.109)	(0.110)	(1.897)	(1.910)
Post × Treatment 2 × Education	0.008	0.008	0.886	0.867	-0.012	-0.012	0.800	0.800
	(0.073)	(0.074)	(2.080)	(2.095)	(0.087)	(0.088)	(2.260)	(2.276)
Baseline controls	No	Yes	No	Yes	No	Yes	No	Yes
R-squared	0.184	0.200	0.107	0.119	0.214	0.228	0.120	0.129

(*continued*)

Table A2.4. (*continued*)

Variables	Results without attrition-weighting				Attrition-weighted results			
	Planted KK15 (dummy)[a]		Share of land under KK15 (%)		Planted KK15 (dummy)		Share of land under KK15 (%)	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel C: Treatment 3 (n=752)								
Post (dummy)	0.012	0.012	0.098	0.098	0.010	0.010	0.081	0.081
	(0.012)	(0.012)	(0.098)	(0.099)	(0.010)	(0.010)	(0.082)	(0.082)
Education (dummy)	-0.000	-0.005	-0.000	-0.111	0.000	-0.006	-0.000	-0.206
	(.)	(0.006)	(0.000)	(0.114)	(.)	(0.008)	(0.000)	(0.182)
Treatment 3 (dummy)	-0.000	-0.012	-0.000	-0.155	0.000	-0.009	-0.000	-0.149
	(0.000)	(0.010)	(.)	(0.185)	(.)	(0.010)	(.)	(0.218)
Post × Education	-0.012	-0.012	-0.098	-0.098	-0.010	-0.010	-0.081	-0.081
	(0.012)	(0.012)	(.)	(0.099)	(0.010)	(0.010)	(0.082)	(0.082)
Post × Treatment 3	0.143***	0.143***	2.702*	2.702*	0.152***	0.152***	3.122*	3.122*
	(0.046)	(0.046)	(1.327)	(1.336)	(0.048)	(0.049)	(1.582)	(1.592)
Post × Treatment 3 × Education	0.115**	0.115**	2.377	2.377	0.126*	0.126*	2.713	2.713
	(0.055)	(0.055)	(1.661)	(1.672)	(0.064)	(0.065)	(2.176)	(2.190)
Baseline controls	No	Yes	No	Yes	No	Yes	No	Yes
R-squared	0.169	0.194	0.105	0.124	0.184	0.210	0.116	0.136

Notes: Coefficient estimates are shown with robust standard errors clustered at farmer group level in parentheses. Post, dummy variable which takes a value of 1 for follow-up round observations (after treatment), and zero for baseline observations. Treatment 1, agricultural training. Treatment 2, agricultural training plus nutrition training. Treatment 3, agricultural training plus nutrition training plus marketing training. Baseline controls include age, gender, education, risk attitude, household size, farm size, value of productive assets, access to credit, distance to market, group official, and county dummy. *, **, and *** denote significance at the 10%, 5%, and 1% level, respectively.

Table A2.5. Effects of Extension Treatments on Technology Adoption, Intent-To-Treat Estimates by Gender

Variables	Results without attrition-weighting				Attrition-weighted results			
	Planted KK15 (dummy)[a]		Share of land under KK15 (%)		Planted KK15 (dummy)		Share of land under KK15 (%)	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: Treatment 1 (n=752)								
Post (dummy)	0.006	0.006	0.051	0.051	0.005	0.005	0.045	0.045
	(0.006)	(0.006)	(0.051)	(0.052)	(0.005)	(0.005)	(0.044)	(0.045)
Male (dummy)	0.000	0.001	0.000	-0.088	0.000	0.004	0.000	-0.208
	(.)	(0.011)	(0.000)	(0.255)	(0.000)	(0.013)	(.)	(0.471)
Treatment 1 (dummy)	0.000	-0.003	0.000	-0.010	-0.000	-0.004	-0.000	-0.139
	(0.000)	(0.007)	(0.000)	(0.168)	(0.000)	(0.010)	(.)	(0.345)
Post × Male	-0.006	-0.006	0.229	-0.051	0.016	-0.005	-0.045	-0.045
	(0.006)	(0.006)	(0.163)	(0.052)	(0.011)	(0.005)	(0.045)	(0.045)
Post × Treatment 1	0.272***	0.272***	5.572***	5.572***	0.271***	0.271***	5.745***	5.745***
	(0.075)	(0.075)	(1.573)	(1.584)	(0.072)	(0.073)	(1.667)	(1.678)
Post × Treatment 1 × Male	-0.141	-0.141	-3.034*	-3.034	-0.133	-0.133	-2.362	-2.362
	(0.114)	(0.114)	(1.744)	(1.756)	(0.127)	(0.128)	(2.564)	(2.581)
Baseline controls	No	Yes	No	Yes	No	Yes	No	Yes
R-squared	0.197	0.208	0.113	0.131	0.178	0.185	0.102	0.117
Panel B: Treatment 2 (n=732)								
Post (dummy)	0.006	0.006	0.051	0.051	0.005	0.005	0.045	0.045
	(0.006)	(0.006)	(0.051)	(0.052)	(0.005)	(0.005)	(0.045)	(0.045)
Male (dummy)	0.000	-0.005	0.000	-0.039	0.000	-0.008	0.000	-0.113
	(0.000)	(0.013)	(.)	(0.300)	(0.000)	(0.013)	(.)	(0.292)
Treatment 2 (dummy)	0.016	0.016	0.328	0.269	0.017	0.017	0.345	0.296
	(0.011)	(0.014)	(0.223)	(0.293)	(0.012)	(0.016)	(0.241)	(0.314)
Post × Male	-0.006	-0.006	-0.051	-0.051	-0.005	-0.005	-0.045	-0.045
	(0.006)	(0.006)	(0.051)	(0.052)	(0.005)	(0.005)	(0.045)	(0.045)
Post × Treatment 2	0.215***	0.215***	4.171***	4.171***	0.247***	0.247***	4.993***	4.993***
	(0.054)	(0.055)	(1.369)	(1.389)	(0.057)	(0.057)	(1.491)	(1.502)
Post × Treatment 2 × Male	0.061	0.061	0.805	0.805	0.037	0.037	-0.398	-0.397
	(0.124)	(0.125)	(2.604)	(2.622)	(0.127)	(0.128)	(2.589)	(2.607)
Baseline controls	No	Yes	No	Yes	No	Yes	No	Yes
R-squared	0.184	0.201	0.107	0.119	0.209	0.228	0.118	0.129

(*continued*)

Table A2.5. (*continued*)

Variables	Results without attrition-weighting				Attrition-weighted results			
	Planted KK15 (dummy)[a]		Share of land under KK15 (%)		Planted KK15 (dummy)		Share of land under KK15 (%)	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel C: Treatment 3 (n=752)								
Post (dummy)	0.006	0.006	0.051	0.051	0.005	0.005	0.045	0.045
	(0.006)	(0.006)	(0.051)	(0.052)	(0.005)	(0.005)	(0.045)	(0.045)
Male (dummy)	0.000	0.003	0.000	-0.633	0.000	0.002	0.000	-0.153
	(0.000)	(0.012)	(0.000)	(0.258)	(.)	(0.013)	(.)	(0.315)
Treatment 3 (dummy)	0.007	0.001	0.058	-0.071	0.008	0.002	0.063	-0.106
	(0.007)	(0.012)	(0.057)	(0.202)	(0.007)	(0.013)	(0.062)	(0.249)
Post × Male	-0.006	-0.006	-0.051	-0.051	-0.005	-0.005	-0.045	-0.045
	(0.006)	(0.006)	(0.051)	(0.052)	(0.005)	(0.005)	(0.045)	(0.045)
Post × Treatment 3	0.204***	0.204***	4.003***	4.003***	0.220***	0.220***	4.901**	4.901**
	(0.055)	(0.056)	(1.378)	(1.389)	(0.062)	(0.062)	(1.791)	(1.802)
Post × Treatment 3 × Male	-0.012	-0.012	-0.508	-0.508	-0.016	-0.016	-1.379	-1.379
	(0.076)	(0.076)	(1.490)	(1.500)	(0.076)	(0.077)	(1.700)	(1.712)
Baseline controls	No	Yes	No	Yes	No	Yes	No	Yes
R-squared	0.154	0.178	0.094	0.114	0.166	0.192	0.107	0.127

Notes: Coefficient estimates are shown with robust standard errors clustered at farmer group level in parentheses. Post, dummy variable which takes a value of 1 for follow-up round observations (after treatment), and zero for baseline observations. Treatment 1, agricultural training. Treatment 2, agricultural training plus nutrition training. Treatment 3, agricultural training plus nutrition training plus marketing training. Baseline controls include age, gender, education, risk attitude, household size, farm size, value of productive assets, access to credit, distance to market, group official, and county dummy. *, **, and *** denote significance at the 10%, 5%, and 1% level, respectively.

3 Agricultural Commercialization and Nutrition in Smallholder Farm Households

Abstract

Commercialization of smallholder agriculture is widely seen as an important route towards rural economic growth. While previous studies have analyzed effects of agricultural commercialization on productivity and income, implications for farm household nutrition have received much less attention. We evaluate the impact of commercialization on household food security and dietary quality, with a special focus on calorie and micronutrient consumption. We also examine transmission channels by looking at the role of income, gender, and possible substitution between consumption of own-produced and purchased foods. The analysis builds on survey data from 805 farm households in Kenya. A control function approach is used to address issues of endogeneity. Generalized propensity scores are employed to estimate continuous treatment effects. Commercialization significantly improves food security and dietary quality in terms of calorie, zinc, and iron consumption. For vitamin A, effects are statistically insignificant. Commercialization contributes to higher incomes and additional nutrients from purchased foods, but does not reduce the consumption of nutrients from own-produced foods, even after controlling for farm size, which can be explained by higher productivity on more commercialized farms. Enhancing market access is important not only for rural economic growth, but also for making smallholder agriculture more nutrition-sensitive.

Key words: agricultural commercialization, market access, continuous treatment effects, nutrition, dietary quality, Africa

JEL *codes*: I15, Q12, Q13, Q18

3.1 Introduction

In spite of global efforts to alleviate hunger and improve nutrition, around 800 million people are still chronically undernourished, and at least 2 billion people suffer from micronutrient deficiencies (FAO 2017; IFPRI 2017). A large proportion of these people are smallholder farmers in developing countries that crucially depend on agriculture as a source of food and income. A key question for improving nutrition is therefore how to make smallholder agriculture more nutrition-sensitive (Smith and Haddad 2015; Pingali and Sunder 2017).[4]

Much of the recent literature on nutrition-sensitive agriculture focuses on the link between on-farm production diversity and farm household diets (Jones, Shrinivas, and Bezner-Kerr 2014; Sibhatu, Krishna, and Qaim 2015; Koppmair, Kassie, and Qaim 2017; Jones 2017). A few of these studies have also pointed at the importance of markets for improving diets, yet capturing farmers' access to markets only in terms of simple proxies such as market distance (Sibhatu et al.2015; Koppmair et al. 2017). Moreover, the dietary indicators that are typically used have limitations. Most studies use household dietary diversity scores, which are suitable for measuring household food security, but not dietary quality (Kennedy, Ballard, and Dop 2013).

Another strand of the literature has analyzed the effects of agricultural commercialization (hereafter referred to as commercialization) on household welfare. But most studies in this direction look at welfare only in terms of income or asset ownership (Tipraqsa and Schreinemachers 2009; Carletto, Kilic, and Kirk 2011; Muriithi and Matz 2015), not nutrition. Commercialization may influence nutrition through various channels, including changes in income, availability of own-produced foods, and gender roles within the farm household (von Braun and Kennedy 1994; Carletto et al. 2015). Income gains can increase the economic access to food, but a substitution of purchased food for own-produced food may also change dietary quality, possibly increasing the consumption of calories but not necessarily micronutrients (Popkin, Adair, and Ng 2012; Remans et al. 2015). Changes in gender roles may occur because men often take stronger control of farm production and income during the process of commercialization (von Braun and Kennedy 1994). And male-controlled income is often spent less on dietary quality and nutrition than female-controlled income (Hoddinott and Haddad 1995, Fischer and Qaim 2012a).

[4] While exact definitions differ, there is broad agreement that nutrition-sensitive agriculture is an approach that seeks to maximize agriculture's contribution to nutrition (IFPRI 2017; Pingali and Sunder 2017).

A few recent studies have analyzed the impact of contract farming or certification programs on household food security in different countries of Africa (Chege, Andersson, and Qaim 2015; Chiputwa and Qaim 2016; Bellemare and Novak 2017). But these studies compare farm households that sell in different marketing channels; no differentiation is made between more and less commercialized households. Very few studies have explicitly analyzed effects of commercialization on nutrition, and those that did have looked at nutrition primarily in terms of calorie consumption and child anthropometrics (von Braun and Kennedy 1994; Carletto, Corral, and Guelfi 2017), not dietary quality.

We add to this literature by analyzing the effects of commercialization on food security and dietary quality, measured in terms of calorie and micronutrient consumption. We estimate average and continuous treatment effects and also analyze transmission channels that were hardly addressed in previous studies (Carletto et al. 2017). The analysis builds on a survey of smallholder farm households in Western Kenya. In Kenya, smallholder farming accounts for 75% of total agricultural output (Olwande et al. 2015). As in most other countries of sub-Saharan Africa, issues of poverty and malnutrition are widespread in the small farm sector (KNBS 2015).

3.2 Conceptual Framework

Figure 3.1 shows a simple conceptual framework that guides our empirical analysis. Commercialization can affect farm household nutrition through various transmission channels. Market sales can reduce the availability of own-produced foods and thus limit consumption through the subsistence pathway. Yet a fall in total food consumption may be prevented through food purchases from the market that are possible through higher cash earnings. Research shows that commercialization is typically associated with income gains through agricultural intensification and use of better technology (von Braun and Kennedy 1994; Muriithi and Matz 2015). Commercialization may also influence the types of crops grown or the livestock species kept on the farm. Closer market integration allows farmers to better harness comparative advantages, so higher levels of specialization are generally expected. A focus on the production of non-food cash crops could further reduce the availability of own-produced foods. Yet, in specific situations, it is also possible that farmers further diversify production, especially when markets for certain niche products that are not traditionally grown for own consumption emerge (Tipraqsa and Schreinemachers 2009).

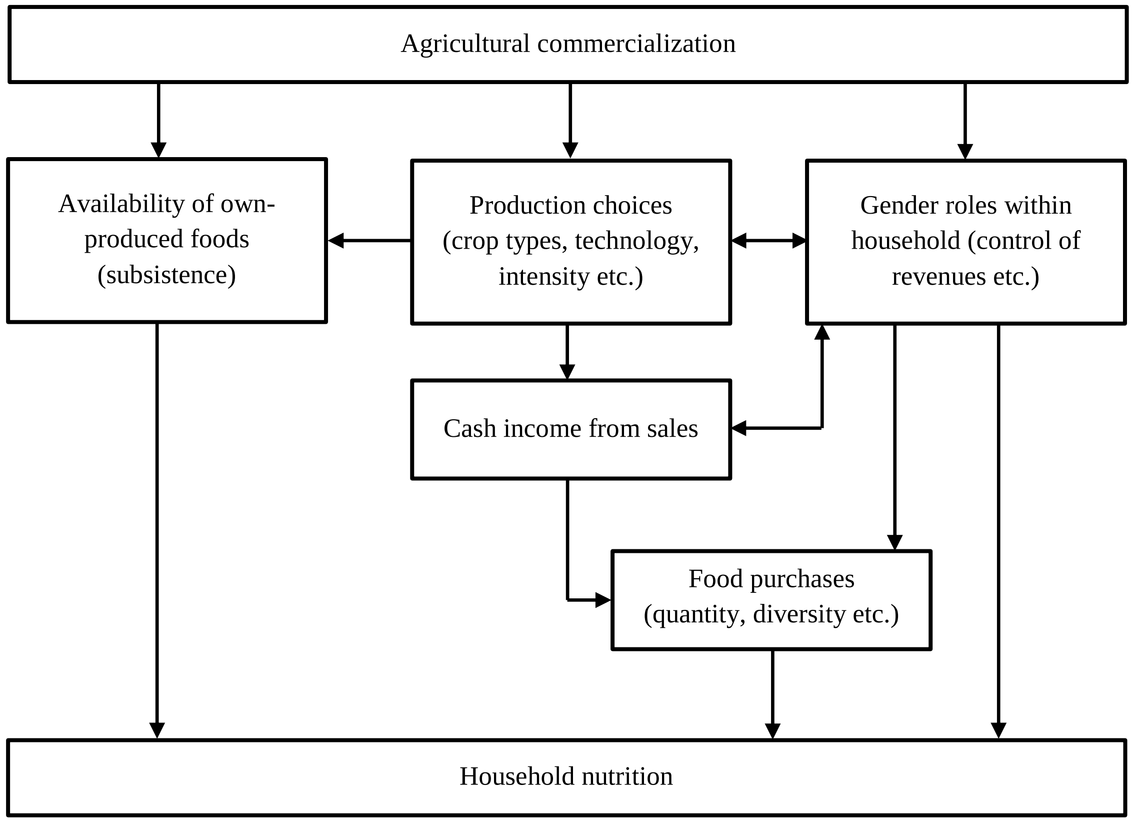

Figure 3.1. Agricultural Commercialization and Household Nutrition Status

Source: Adapted from von Braun and Kennedy (1994) and Chege et al. (2015).

In an African context, levels of commercialization, types of crops grown, and technologies used can also have important effects on gender roles within the farm household. This requires deviation from the unitary household model that assumes that a single agent makes all production and consumption decisions (Haddad, Hoddinott, and Alderman 1997). Subsistence food crops are often produced and controlled by women, whereas crops that are primarily produced to generate cash income are typically controlled by men (von Braun et al. 1994; Fischer and Qaim 2012a). Research shows that female-controlled income is often particularly beneficial for household nutrition, as women tend to spend more on food, dietary quality, and healthcare than men (Hoddinott and Haddad 1995; Chege et al. 2015). Hence, commercialization may possibly have a negative partial effect on household nutrition through this gender pathway, even though effects may vary depending on the empirical context.

To better understand the role of the different transmission channels and the overall effect of commercialization on farm household nutrition, crucial questions are to what extent own-produced food is replaced by purchased food, and whether this shift makes diets more or less nutritious from a calorie and micronutrient perspective. It is often assumed that the subsistence pathway is particularly important for dietary quality, because purchased food may be more processed and less nutritious (Remans et al. 2015; Jones 2017), even though the evidence base for this assumption is relatively thin.[5] We will analyze these questions explicitly in the empirical analysis below.

The conceptual framework builds on a common assumption in many farm household models, namely that production and consumption decisions are non-separable, due to price risks, transaction costs, and imperfect factor and product markets (Key, Sadoulet, and de Janvry 2000). Under this assumption, production decisions affect consumption decisions, and vice versa, and ultimately market supply and demand. We also assume imperfect substitutability between own-produced and purchased foods and imperfect knowledge about the nutritional value of different food items. These assumptions are reasonable for the rural small farm context, where market failures are commonplace.

3.3 Estimation Strategy

3.3.1 Basic Model

We start the analysis by estimating the overall effect of commercialization on nutrition with regression models of the following type:

$$N_i = \alpha_0 + \alpha_1 C_i + \alpha_2 \boldsymbol{X}_{i1} + \varepsilon_{i1}, \tag{5}$$

where N_i is the nutrition indicator for household *i*. We use different nutrition indicators, namely calorie and micronutrient – vitamin A, zinc, and iron – consumption levels, and estimate separate regressions for all of them. Details of these indicators are described further below. C_i in equation (5) is the level of commercialization, $\boldsymbol{X}_i$ is a vector of control

[5] The assumption of purchased foods being less healthy than own-produced foods is partly related to the growth of supermarkets and other modern retail outlets in many developing countries. These modern retailers often focus on the sales of highly processed, energy-dense foods that are rich in fats and sugars (Popkin et al. 2012; Tschirley et al. 2015). However, in sub-Saharan Africa, modern retailers are primarily found in larger urban centers up till now. Rural farm households buy most of their food in traditional local markets and shops. In fact, a recent study with data from rural Ethiopia showed that purchased food contributes more to dietary quality than own-produced food (Sibhatu and Qaim 2017). Fresh fruits and vegetables are often not available from own production all the year around. And even processed foods are not always bad for dietary quality, for instance when micronutrients are added through industrial fortification.

variables, and ε_{i1} is a random error term. The level of commercialization (C_i) is defined as the proportion of farm output sold at the market. Control variables ($\boldsymbol{X}_i$) include age, gender, and education of the household head, as well as other farm, household, and contextual variables that may affect diets and nutrition.

In this model, we are particularly interested in the treatment effect α_1. Positive estimates for α_1 would mean that commercialization contributes to improved nutrition, while negative estimates would mean that commercialization contributes to poorer nutrition. It is possible that the sign of α_1 differs between the nutrition indicators. For instance, if it is true that households substitute energy-dense purchased foods for more nutritious own-produced foods we would expect a positive coefficient α_1 in the calorie consumption model and possibly negative coefficients in the micronutrient consumption models.

3.3.2 Addressing Possible Issues of Endogeneity

If $\boldsymbol{X}_i$ in equation (5) includes all the factors that influence commercialization, and there is no correlation between C_i and ε_{i1}, then ordinary least squares (OLS) would produce unbiased estimates of α_1. However, it is possible that there are unobserved factors that jointly influence C_i and N_i, which would lead to endogeneity bias. For instance, unobserved heterogeneity could occur through differences in farmers' ability or entrepreneurial skills, which are difficult to measure in household surveys. Furthermore, there could be issues of reverse causality, where better nutrition would make farmers more productive, thus possibly contributing to higher levels of commercialization. Finally, measurement error could be a cause of endogeneity, even though we put substantial effort into the collection of high-quality data.

We account for these potential sources of endogeneity through a control function (CF) approach (Smith and Blundell 1986; Rivers and Voung 1988; Wooldridge 2015). An alternative would have been standard instrumental variables (IV) estimators. We prefer the CF approach, because the treatment variable (C_i) is bounded between zero and one (and has nontrivial zero observations), leading to non-linear corner solution. In non-linear models, the CF approach was shown to be more efficient than standard IV estimators (Wooldridge 2015; Verkaart et al. 2017). Generally, the CF approach also easily estimates average treatment effects, which are potentially more policy-relevant than the local average treatment effects that are typically estimated with standard IV estimators (Wooldridge 2015).

The CF approach entails predicting residuals from a first-stage model of the determinants of commercialization, and including the predicted residual term as an additional regressor (a control function) in the nutrition outcome model in equation (5). This CF approach requires at least one valid instrument in the first-stage regression. A statistically significant coefficient of the predicted residual term in equation (5) would imply that commercialization is endogenous to the nutrition indicators and would also correct for the resulting bias. An insignificant residual term would fail to reject the null hypothesis of exogeneity of C_i. In that case, OLS would be preferred. Since C_i is bounded between zero and one, we estimate the first-stage regression using a generalized linear model (GLM) with a binomial family and a logit link. This is important to obtain consistent residual predictions for use in the second-stage regression (Papke and Wooldridge 1996).

3.3.3 Choice of Instruments

A valid instrument must be strongly correlated with commercialization (instrument relevance), but uncorrelated with omitted variables that may affect nutrition (instrument exogeneity), except indirectly through commercialization (Imbens and Wooldridge 2009). We were able to identify two instruments for commercialization that fulfill all the requirements of validity. The first instrument is the average number of motorcycles owned by households living in the same ward, and the second instrument is the average number of main market sellers in the ward. A ward is an administrative unit in Kenya that is larger than a village, but smaller than a sub-county. As explained below, our survey covered 26 wards in two counties and 8 sub-counties. The average number of households surveyed in each ward is 31. We explain the two instruments and related tests in the following.

The first instrument – number of motorcycles in the ward – was constructed by counting the number of motorcycles owned by sample households in each ward (excluding the farmer of interest himself/herself) and then dividing by the number of households to obtain an average.[6] Over 90% of the farmers in our sample do not own any motorized means of transportation (the average number of motorcycles owned by households is 0.08). Nevertheless, the distance to the market is often too far to make significant sales without a motor vehicle. As most of the local roads are not paved and public transport services hardly exist, owners of motorcycles tend to offer transport services also for other households living in the same area. Farmers

[6] In our survey, we sampled from households that are organized in farmer groups, so the instrument represents the average motorcycle ownership of households that are organized in farmer groups, not necessarily of all households living in the ward. Details of the sampling framework are provided below.

themselves often use these services, as well as local traders who buy at the farm gate and sell in the marketplace. Hence, more motorcycles in the ward mean better market access. Indeed, the average number of motorcycles in the ward is significantly correlated with the degree of commercialization (p-value=0.018; table A3.1 in the appendix), which is the condition for instrument relevance.

As we use the average number of motorcycles owned by households in the ward, rather than individual ownership, this instrument is not directly correlated with any of the household nutrition variables used in this analysis. Nor is the instrument significant when included as an additional regressor in equation (5). Results of these tests are shown in table A3.2 in the appendix. Nevertheless, one could imagine that the number of motorcycles could also be a proxy for higher levels of wealth and income in the ward. Moreover, motorcycles could also be used to transport inputs or extension officers who deliver agricultural, health, and nutrition training to local farm households, which might influence nutrition indirectly through various hidden channels. To test for these possibilities, we correlated the instrument with several indicators of living standard and wealth at the ward level, such as average education, household income, farm size, and productive assets (table A3.3). We also correlated the instrument with household-level nutritional knowledge scores and the use of various purchased farm inputs, such as seeds, fertilizers, and pesticides (table A3.4). All of the correlation coefficients are statistically insignificant, so that the conditions for instrument exogeneity seem to be fulfilled.

The second instrument – average number of main market sellers in the ward – was constructed by counting the number of sample households in each ward that sell at least some of their produce in the main agricultural market (excluding the farmer of interest himself/herself). This number was divided by the number of households surveyed in the ward. The main agricultural market in each ward is larger than the village markets and is typically the most important point of sale for larger quantities of farm output (only tea and coffee are sold at special collection centers). In our sample, 32% of the households sell at least part of their produce in the main market, and these households tend to be more commercially-oriented (table A3.5). Many of the other households sell smaller quantities of farm output in the village markets or to traders at the farm gate.

Using the number of main market sellers in the ward as an instrument for own commercialization is motivated by the recent literature on the role of neighborhood effects and social networks for farmers' technology adoption and marketing decisions (Krishnan and Patnam 2013; Andersson et al. 2015; Magnan et al. 2015). Social networks in the neighborhood can not only improve the flow of information but can also facilitate the coordination of joint transport and marketing activities (Andersson et al. 2015). Indeed, the number of main market sellers in the ward is highly correlated with the own level of commercialization (*p*-value=0.000, table A3.1), thus satisfying the instrument relevance condition. The instrument is also uncorrelated with all the nutrition indicators (table A3.2).

One could argue that farm households selling in the main markets may cluster in certain locations that are richer and more developed than others. However, farm households in rural Kenya do not actively decide about their location, as they live on their land, which is usually inherited from one generation to the next. We tested whether the instrument is correlated with mean wealth characteristics at the ward level, but none of the correlation coefficients was found to be significant (table A3.3). Nor did we find significant correlation between the number of main market sellers and individual input use or nutrition knowledge (table A3.4), making it unlikely that the instrument would affect nutrition through channels other than commercialization.

Using both instruments, we also formally tested for over-identification. Based on the test results, which are shown in table A3.1 in the appendix, we could not reject the null hypothesis of instrument exogeneity. We acknowledge that completely eliminating all possible issues of endogeneity is always challenging with cross-section observational data, but based on the various tests the instruments seem to be valid, so that cautious causal inference should be in order.

3.3.4 Analyzing Transmission Channels

As discussed above, important questions to better understand the transmission channels between commercialization and nutrition are to what extent purchased foods are substituted for own-produced foods and how this affects dietary quality. To analyze this in more detail, we re-estimate the models in equation (5), but this time differentiating between calories and micronutrients from purchased and own-produced foods. If households primarily purchased energy-dense foods in the market, we would expect a positive effect of commercialization on calorie consumption, but not micronutrient consumption from purchased foods. The effects of commercialization on calorie and micronutrient consumption from own-produced foods will depend on possible changes in farm productivity and production diversity. For these models, we also use the control function approach with the same two instruments.

Beyond analyzing possible substitution effects in household food sources, we are also interested in better understanding the possible role of income and gender pathways. We model these pathways explicitly with the following equations:

$$N_i = \beta_0 + \beta_1 Y_i + \beta_2 G_i + \beta_3 \boldsymbol{X}_i + \varepsilon_{i2}, \quad (6)$$

$$Y_i = \delta_0 + \delta_1 C_i + \delta_2 \boldsymbol{X}_i + \varepsilon_{i3}, \quad (7)$$

$$G_i = \gamma_0 + \gamma_1 C_i + \gamma_2 \boldsymbol{X}_i + \varepsilon_{i4}. \quad (8)$$

Equation (6) models nutrition (N_i) as a function of household income (Y_i) and gender roles within the household (G_i), measured in terms of a dummy that takes a value of one if a male household member controls the farm revenues, and zero otherwise. Given the discussion in the conceptual framework section, we would expect a positive coefficient estimate for β_1 and a negative estimate for β_2. In equations (7) and (8), income and gender roles are considered endogenous and modeled as functions of commercialization (C_i). We would expect positive coefficient estimates for δ_1 and γ_1, meaning that commercialization increases household income and the likelihood of male control of farm revenues. In all three equations we control for other socioeconomic variables ($\boldsymbol{X}_i$). We estimate equations (6) and (7) with OLS, and equation (8) with linear probability and probit estimators. The results of these models will be interpreted cautiously in terms of associations, not causality, because of the endogeneity of Y_i and G_i.

3.3.5 Continuous Treatment Effects

The models in equations (5) to (8) help to establish the average treatment effects of commercialization on nutrition and the underlying transmission channels. But commercialization is a continuous treatment variable, and it is possible that the effects vary by level of commercialization. For instance, when a subsistence farm starts to make the first market sales, the marginal effects on calorie and micronutrient consumption may be higher or lower than when a farm that already sells much of its produce further increases the level of commercialization. To account for possible non-linearity, we use the generalized propensity score (GPS) approach to estimate continuous treatment effects of commercialization (Hirano and Imbens 2004; Kluve et al. 2012; Guardabascio and Ventura 2014). The GPS method controls for observed heterogeneity between households with different treatment exposure, but not for possible unobserved heterogeneity.

The GPS approach involves three stages (Kassie, Jaleta, and Mattei 2014). First, the generalized propensity scores are generated based on observed covariates. Given the nature of C, we estimate the GPS using GLM with a binomial family and a logit link (fractional logit). The first stage also involves testing covariate balancing properties. Second, the conditional expected values of the outcome variables (nutrition indicators) are estimated as a function of treatment exposure (level of commercialization) and the GPS. For these estimates, we use a flexible function with quadratic approximation. Given that the nutrition indicators are continuous variables, these models are estimated with OLS. Third, the average dose-response function is estimated. The dose-response function depicts for every treatment exposure level the direction and magnitude of the causal relationship between commercialization and nutrition, after controlling for any observed covariate bias (Hirano and Imbens 2004).

We estimate the dose-response function by averaging the expected nutrition outcome at each level of commercialization (C) as follows:

$$E\left[\hat{N}_i(C)\right] = \frac{1}{n}\sum_{i=1}^{n}\left[\hat{\alpha}_0 + \hat{\alpha}_1 C + \hat{\alpha}_2 C^2 + \hat{\alpha}_3 \hat{r}(C, \boldsymbol{X}_i) + \hat{\alpha}_4 \hat{r}(C, \boldsymbol{X}_i)^2 + \hat{\alpha}_5 C\hat{r}(C, \boldsymbol{X}_i)\right], \qquad (9)$$

where n is the number of observations, the $\hat{\alpha}$ values are parameters estimated at the second stage, and $\hat{r}(C, \boldsymbol{X}_i)$ is the predicted value of the conditional density of treatment at varying levels of commercialization. Results of the dose-response functions are presented graphically.

3.4 Data and Variable Measurement

3.4.1 Farm Household Survey

This study builds on data collected through a survey of smallholder farm households in Kisii and Nyamira counties in Western Kenya between October and December 2015. Given the high population density in the study area, farms are very small with farm sizes ranging between 0.5 and 5 acres (Kisii County Government 2013; Nyamira County Government 2013). Like many other parts of sub-Saharan Africa, agriculture in the study area is largely semi-subsistent. Higher levels of commercialization are curbed primarily by high transport and transaction costs related to poor infrastructure and other market failures. In terms of nutritional indicators, Kisii and Nyamira are similar to the national average in Kenya (KNBS 2015). The prevalence of child stunting, the most common anthropometric measure of child undernutrition, is around 26% in both counties (KNBS 2015).

A recent census of farm households in Kisii and Nyamira was not available. However, many farmers are organized in farmer groups that are registered with the Ministry of Gender, Children, and Social Development. We therefore decided to cluster our survey by farmer groups. Building on Ministry registries and with support from Africa Harvest, a non-governmental organization active in the region, a list of all existing farmer groups in Kisii and Nyamira was constructed. From this list, we excluded a few groups that had received specific development support during the last two years.[7] From the remaining groups, we randomly selected 48 for inclusion in the survey. These groups varied in size, most of them had around 20-30 members. Prior to the survey, we updated group membership lists together with the group leaders. Depending on group size, we randomly selected 15-20 member households from each group, resulting in a total sample size of 824 farm households, distributed over 8 different sub-counties and 26 wards.

Data from sample households were collected through face-to-face interviews carried out in the local language with the household head and/or the spouse.[8] A carefully designed and pretested questionnaire was used, capturing details on household demographics, agricultural production and marketing, other economic activities of the household, food and non-food consumption and expenditures, and contextual characteristics. All details about agricultural production and marketing were captured for a period of 12 months. For the food consumption data we used shorter recall periods, as explained in more detail below. For a few of the sample households, relevant variables are missing. The analysis is carried out with observations from 805 households for which complete data are available.

3.4.2 Measuring Nutrition

To assess the effects of commercialization on household nutrition, we need appropriate nutrition indicators. There are various ways to assess nutrition at individual and household level, including clinical measures, anthropometric measures, and food consumption-based measures, among others (de Haen, Klasen, and Qaim 2011; Masset et al. 2012; IFPRI 2017). Clinical and anthropometric measures are the most precise indicators of individual nutrition

[7] Excluded groups had received agricultural trainings for the adoption and use of tissue culture banana technology. These groups were located in specific sub-counties, geographically separated from the rest, so that spillover effects are unlikely.

[8] Of the respondents, 60% were female and 40% were male.

status, but they are less suitable to assess details of people's food sources and dietary quality, which is the focus of our study. Hence, we use food consumption data, from which we calculate various measures of food security and dietary quality.

The survey questionnaire included a food consumption recall, capturing the quantity of more than 130 different food items consumed by all household members over a period of 7 days. Survey respondents were also asked to specify the source of each food item consumed, including market purchases, own production, gifts, and other sources. To increase data accuracy, this part of the questionnaire was carried out with the person responsible for food preparation in the household. Based on the food quantities consumed, we calculated edible portions, which were then converted to calorie and micronutrient levels using food composition tables for Kenya (Sehmi 1993). For individual food items not included in these tables, we used food composition tables for neighboring Tanzania (Lukmanji et al. 2008). In terms of micronutrients, we focus on vitamin A, zinc, and iron. Deficiencies in vitamin A, zinc, and iron pose serious health challenges in many developing countries, so that consumption levels of these three micronutrients are considered important proxies of healthy diets and nutrition (IFPRI 2017).

We divided calorie and micronutrient consumption at household level by adult male equivalents (AE) to make the values comparable across households of different size, age and sex composition, and levels of physical activity (Chege et al. 2015; Chiputwa and Qaim 2016; Coates et al. 2017). These consumption values per AE are the nutrition indicators (N_i) used as outcome variables in the econometric models. For the descriptive analysis, we calculate a few additional indicators to further illustrate the local nutrition situation. We use minimum consumption thresholds to characterize undersupplied households (FAO, WHO, and UNU 2001; IOM 2006; IZiNCG 2004). A household is considered undernourished when it consumes less than 2400 kcal per AE and day. A household is deficient in vitamin A when it consumes less than 625 µg of retinol equivalents (RE). For zinc and iron, the thresholds are 15.0 mg and 18.3 mg, respectively.

For the descriptive analysis, we also construct two simpler nutrition-related indicators that were used in the recent literature on linkages between farm production and dietary diversity (Jones et al. 2014; Sibhatu et al. 2015; Koppmair et al. 2017). First, we compute the household dietary diversity score (HDDS), which counts the number of food groups consumed during the 7-day recall period. The maximum number of food groups in the HDDS is 12. The higher the value, the better is the household's general access to food (Kennedy, Ballard, and Dop 2013). Second, we compute a different dietary diversity score with a maximum of 10 food groups (DDS10). The DDS10 places higher emphasis on food groups that are important from a micronutrient perspective and is therefore a better proxy of dietary quality (FAO 2016).

Using household-level food consumption data from a 7-day recall to assess diets and nutrition has become common in the food economics literature (de Haen et al. 2011; Jones et al. 2014; Zezza et al. 2017), even though this approach also has its drawbacks. First, it measures food availability, not actual food intake. Second, possible issues of intra-household distribution are not accounted for. Fortunately, for a subsample of the households we also collected individual-level nutrition data for adults and children, including a 24-hour dietary recall and anthropometric measurements. Table A3.6 in the appendix shows significantly positive correlation coefficients between the household-level and individual-level measures, suggesting that the household-level indicators are suitable proxies of individual diets and nutrition. This is in line with recent research in other geographical contexts (Coates et al. 2017; Sununtnasuk and Fiedler 2017). That we use household-level data in spite of the availability of individual-level data is due to the relatively small sample for children in particular (n=224). Moreover, the 24-hour recall was conducted only once, so that day-to-day variation in diets could not be captured. This is less of an issue when using 7-day recall data.

In addition to day-to-day variation, diets and nutrition often vary seasonally following the agricultural production cycle (Sibhatu and Qaim 2017). Such seasonal variation is not fully captured in our cross-section data. However, in the study area seasonal variation seems to be relatively small. In Kisii and Nyamira counties, farmers typically have two production seasons per year. But, due to abundant rainfall, some cropping also occurs outside these regular seasons, so real lean months are hardly discernible (Kisii County Government 2013; Nyamira County Government 2013). Our survey was conducted between October and December 2015, with some of the households interviewed earlier and others later during this period. Since harvesting of the regular long-rain season is typically completed in the month of September, calorie and nutrient consumption could possibly be higher in October and November, the two months directly following the harvest, than in December. We did not find significant correlations between a "December survey" dummy variable and the nutrition indicators (table A3.7). Nevertheless, to avoid any possible bias due to seasonality, we use the "December survey" dummy as an additional control variable in the regression models.

3.4.3 Measuring Commercialization

While 97% of the households in our sample sell some of their farm produce, more than half of the farm output (56%) is kept for home consumption. This underlines that the level of commercialization is limited. As mentioned above, 32% of the households sell in the agricultural main market. Around 50% sell in local village markets, and 73% sell at least some of their harvest at the farm gate (figure A3.1 in the appendix). Contract farming is rare in the study region for typical food crops, but exists for cash crops such as tea and coffee.[9] These cash crops have to be delivered to special collection centers.

Following von Braun and Kennedy (1994) and Carletto et al. (2017), we construct a commercialization index defined as the share of the total value of farm output sold (value of output sold divided by value of total farm output) during the 12-months period covered by the survey. This includes both crop and livestock enterprises, regardless of where the products were sold (farm gate, main market, collection center etc.). The commercialization index is a continuous variable ranging between zero and one.[10] For the construction of the index, price data are required to value the quantities of farm output. Prices may vary, even for identical commodities, and they are not observed for all households. For better comparison, we use average sales prices reported by sample households to value farm output.

For robustness check, we also compute four other measures of commercialization. First, we use a maize commercialization index, calculated as the share of total maize production sold during the last 12 months (in quantity terms). Maize is the most important staple food in Kenya and is widely grown by sample households primarily for home consumption. Second, we use a crop commercialization index, computed as the share of total crop output sold (in value terms). Third, we use a livestock commercialization index, computed as the share of total livestock products sold (in value terms). Finally, we re-calculate the commercialization index for all crop and livestock products as defined above, but using ward-level prices instead of total sample mean prices to value farm output.

In addition to running the regressions with these alternative measures of commercialization, we correlated the commercialization index with these alternative measures and several others that describe different facets of commercialization, such as the absolute value of product

[9] While we collected data on the point of sale and sales revenues of all farm commodities, details about contracts were not covered in the survey. However, we know that the contracts between tea and coffee companies or cooperatives and smallholder farmers are pure marketing contracts that do not include the provision of inputs.

[10] This output-side index captures households' revealed marketing behavior (Carletto et al. 2017). As there are only very few farm households that sell zero farm output, a continuous index is better suited for the analysis than a binary commercialization variable.

sales, the share of the land under cash crops, or the value of farm inputs used. All of these measures are positively correlated with our commercialization index at high levels of statistical significance (table A3.8).

3.5 Descriptive Statistics

3.5.1 Socioeconomic Characteristics

Table 3.1 presents selected summary statistics for the full sample, as well as differentiated by level of commercialization (additional variables are shown in table A3.9 in the appendix). For the descriptive part, we subdivide the sample into commercialization quartiles and compare the 25% most commercialized households (MC25%) with the 25% least commercialized households (LC25%).

The average household sells 44% of its total farm output. This share ranges between 70% for the most commercialized and 16% for the least commercialized households. The most commercialized households sell more in all types of markets than the least commercialized ones, including sales at the farm gate, but the largest differences are observed for sales in the main agricultural markets (figures A3.1 and A3.2), as one would expect. The level of commercialization is positively associated with farm size, education, household income, and several other socioeconomic variables (table 3.1). More commercialized farms also use more inputs, such as fertilizers and pesticides, and they have significantly higher land productivity.

Sample farms are highly diversified, producing around 13 different crop and livestock species on average. Such a high level of farm diversity is typical for many regions in Africa (Sibhatu et al. 2015). Sample farms produce a number of different food crops, such as maize, beans, sweet potatoes, bananas, and different types of leafy vegetables. Many also keep chicken, sheep, goats, and sometimes cattle. In terms of cash crops, tea, coffee, and sugarcane are grown by many farmers. Strikingly, more commercialized households are more diversified than less commercialized households, suggesting that under the given conditions commercialization does not lead to higher levels of farm specialization.

Table 3.1. Summary Statistics by Level of Commercialization

Variables	Total mean	MC25% mean	LC25% mean	Mean difference
Socioeconomic characteristics				
Commercialization (share of farm output sold, 0-1)	0.44	0.70	0.16	0.55***
	(0.21)	(0.09)	(0.09)	
Male household head (dummy)	0.77	0.82	0.67	0.15***
	(0.42)	(0.39)	(0.47)	
Education of household head (years)	8.94	9.69	7.80	1.89***
	(3.77)	(3.19)	(4.09)	
Farm size (acres)	1.61	2.04	1.14	0.90***
	(1.27)	(1.55)	(0.95)	
Household income (1,000 Ksh/year)	180.53	281.36	90.69	190.67***
	(218.46)	(285.81)	(103.12)	
Farm productivity and input use				
Value of crop output (1,000 Ksh/acre)	75.81	105.13	70.32	34.80***
	(81.94)	(110.42)	(97.12)	
Fertilizer cost (1,000 Ksh/acre)	6.27	6.57	5.38	1.19**
	(5.48)	(6.34)	(4.52)	
Pesticide cost (1,000 Ksh/acre)	0.66	0.91	0.33	0.58***
	(1.63)	(2.04)	(1.08)	
Production diversity (no. of crop/livestock species)	12.87	13.12	11.68	1.45***
	(4.66)	(4.94)	(4.40)	
Observations	805	201	202	403

Note: Standard deviations are shown in parentheses. MC25%, 25% most commercialized households; LC25%, 25% least commercialized households; Ksh, Kenyan shillings (1 US dollar = 96.3 Ksh). *, **, and *** significant at 10%, 5%, and 1% level, respectively. Additional variables are shown in table A3.9 in the appendix.

3.5.2 Nutrition Indicators

Table 3.2 shows summary statistics for the nutrition indicators. Around 27% of the sample households are undernourished (calorie-deficient). Even higher proportions are deficient in zinc, iron, and vitamin A, pointing at sizeable nutritional problems. More commercialized households consume significantly higher amounts of calories and micronutrients. Thus, they are also less affected by nutritional deficiencies than less commercialized households. Only for vitamin A deficiency, the difference is not statistically significant. The dietary diversity scores also suggest better access to food and higher dietary quality among more commercialized households.

Table 3.2. Summary Statistics of Nutrition Indicators by Level of Commercialization

Variables	Total mean	MC25% mean	LC25% mean	Mean difference
Total calorie consumption (kcal/day/AE)	3286.06	3584.42	2973.07	611.35***
	(1273.73)	(1294.94)	(1065.46)	
Prevalence of undernourishment (%)	26.58	17.91	33.66	-15.75***
	(44.21)	(38.44)	(47.37)	
Total vitamin A consumption (µg RE/day/AE)	1242.55	1406.09	1140.09	266.01*
	(1393.24)	(1542.6)	(1231.14)	
Prevalence of vitamin A deficiency (%)	36.65	33.33	37.62	-4.29
	(48.21)	(47.26)	(48.56)	
Total zinc consumption (mg/day/AE)	19.67	21.07	18.25	2.82***
	(8.70)	(8.72)	(7.67)	
Prevalence of zinc deficiency (%)	32.42	24.38	40.10	-15.72***
	(46.84)	(43.04)	(49.13)	
Total iron consumption (mg/day/AE)	22.10	25.04	18.61	6.43***
	(13.31)	(15.21)	(9.76)	
Prevalence of iron deficiency (%)	47.20	40.30	56.93	-16.63***
	(49.95)	(49.17)	(49.64)	
Household dietary diversity score (HDDS)	9.42	9.57	9.03	0.54***
	(1.44)	(1.34)	(1.58)	
Dietary diversity score, 10 food groups (DDS10)	7.07	7.24	6.76	0.48***
	(1.46)	(1.35)	(1.67)	
Observations	805	201	202	403

Note: Standard deviations are shown in parentheses. MC25%, 25% most commercialized households; LC25%, 25% least commercialized households; AE, male adult equivalent; RE, retinol equivalent. *, **, and *** significant at 10%, 5%, and 1% level, respectively.

Figure 3.2 shows a breakdown of the sources of calorie and micronutrient consumption. For calorie, zinc, and iron, market purchases are as important as, or even more important than own production. This is true even for the least commercialized households. Interestingly, for more commercialized households the role of own production for diets does not decrease.

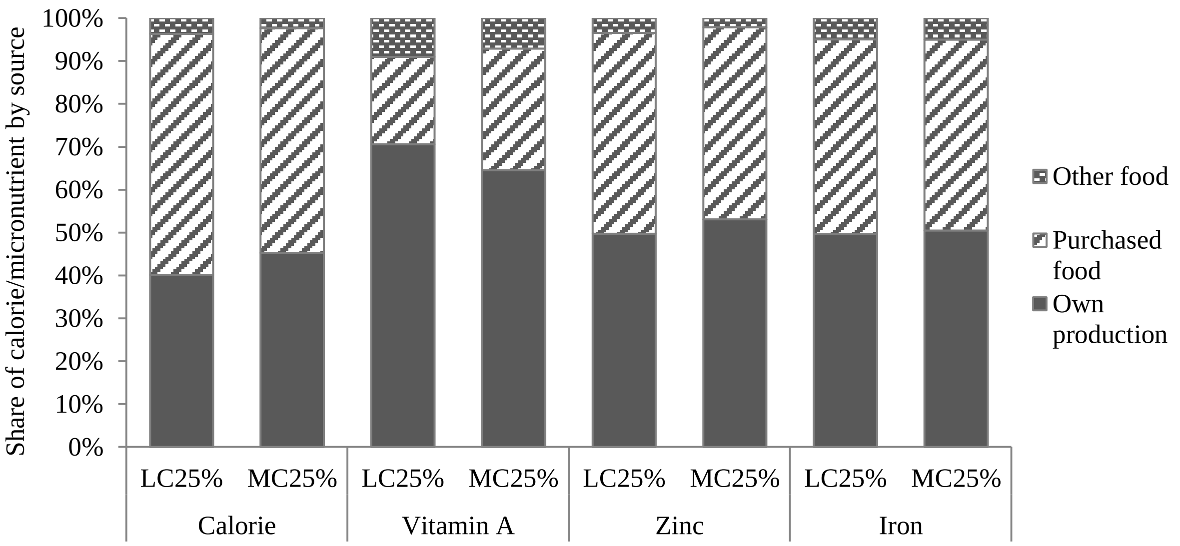

Figure 3.2. Share of Calorie and Nutrient Consumption from Different Sources

Note: LC25%, 25% least commercialized households; MC25%, 25% most commercialized households.

This is a first indication that the cash income generated through output sales may be used to buy additional food, rather than replacing own-produced food. Higher productivity on more commercialized farms allows larger market sales without reducing home consumption. Only for vitamin A, the situation is somewhat different. Own production plays the dominant role for vitamin A, especially in the least commercialized households. Tables A3.10-A3.12 in the appendix show further details of which food groups are particularly important for micronutrient consumption from market and own-produced sources.[11]

3.6 Econometric Results

3.6.1 Endogeneity Tests

We start the discussion of the estimation results by looking at the Hausman tests for endogeneity of commercialization. As explained, we use a CF approach with the average number of motorcycles and the average number of main market sellers in the ward as two valid instruments. The coefficient estimates for the residual terms included in the second-stage equations are shown in table 3.3, for the total calorie and micronutrient consumption models, as well as for the models that distinguish between the consumption of purchased and own-produced foods.

Table 3.3. Testing for Endogeneity of Commercialization Using Control Function Approach

Nutrition indicators	Coefficient	*p*-value	Conclusion
Total calories consumed (kcal/day/AE)	-119.350 (242.578)	0.623	Exogenous
Calories from purchased food (kcal/day/AE)	-107.727 (201.177)	0.592	Exogenous
Calories from own-produced food (kcal/day/AE)	-135.303 (221.409)	0.541	Exogenous
Total vitamin A consumed (µg RE/day/AE)	-444.401 (287.197)	0.122	Exogenous
Vitamin A from purchased food (µg RE/day/AE)	-3.268 (98.334)	0.973	Exogenous
Vitamin A from own-produced food (µg RE/day/AE)	-405.103 (227.991)	0.076	Endogenous
Total zinc consumed (mg/day/AE)	-2.690 (1.817)	0.139	Exogenous
Zinc from purchased food (mg/day/AE)	-0.697 (1.449)	0.631	Exogenous
Zinc from own-produced food (mg/day/AE)	-1.718 (1.617)	0.288	Exogenous
Total iron consumed (mg/day/AE)	1.435 (2.923)	0.624	Exogenous
Iron from purchased food (mg/day/AE)	-0.547 (2.062)	0.791	Exogenous
Iron from own-produced food (mg/day/AE)	1.651 (2.137)	0.440	Exogenous

Note: Coefficients of the residual terms included in the model in equation (5) are shown with bootstrapped standard errors in parentheses.

[11] While we did not collect data on where exactly each individual food item was purchased, traditional village and ward markets (often the same markets where farm outputs are also sold) are the major sources of food purchases for sample households. Supermarkets exist in the county capitals, but not yet in rural areas of Kisii and Nyamira.

In all models, except for own-produced vitamin A, the residual-terms are statistically insignificant. Hence, we cannot reject the null hypothesis that commercialization is exogenous. Based on these Hausman test results, we proceed with OLS, which is more efficient in this case. However, since failure to reject the null hypothesis of the Hausman test is rarely a convincing proof of exogeneity, we also show results of the CF models in the appendix (tables A3.1, A3.13, and A3.14). OLS and CF model results are similar and support the same conclusions.

3.6.2 Basic Model Results

The estimation results of the basic model with total calorie and micronutrient consumption levels as dependent variables are shown in table 3.4. Commercialization has positive and significant effects on all nutrition indicators, except for vitamin A.[12] The commercialization index ranges between zero and one, meaning that a 10 percentage point increase in the level of commercialization increases the consumption of calories by 68.0 kcal (about 3% of the minimum consumption threshold), of zinc by 0.34 mg (2%), and of iron by 0.55 mg (3%) per AE and day. These effects support the hypothesis that commercialization improves farm household nutrition.

[12] In this article, we analyze the effects of commercialization on 12 nutrition indicators. Multiple hypotheses testing can increase the probability of type I errors (detecting false positives). To control for type I errors with multiple hypotheses, we calculate adjusted *p*-values (a.k.a *q*-values) using the two-stage procedure proposed by Benjamini, Krieger, and Yekutieli (2006), following the steps outlined by Anderson (2008). These *q*-values for the treatment effect estimates are shown in square brackets in tables 3.4-3.6.

Table 3.4. Commercialization Effects on Total Calorie and Nutrient Consumption

Variables	Calories (kcal/day/AE)	Vitamin A (µg RE/day/AE)	Zinc (mg/day/AE)	Iron (mg/day/AE)
Commercialization (0-1)	679.945***	136.876	3.447**	5.546**
	(204.332)	(227.202)	(1.307)	(2.220)
	[0.025]	[0.337]	[0.025]	[0.025]
Age of household head (years)	5.532*	4.890	-0.019	0.000
	(2.891)	(5.447)	(0.024)	(0.049)
Age squared (years)	0.279	-0.235	0.001	-0.001
	(0.216)	(0.282)	(0.002)	(0.003)
Male household head (dummy)	-67.539	108.732	-3.269***	0.778
	(102.928)	(109.966)	(0.761)	(1.116)
Education of household head (years)	37.694***	14.243	0.199**	0.489***
	(11.624)	(14.934)	(0.088)	(0.161)
Household size (AE)	-364.051***	-100.929***	-1.783***	-1.858***
	(31.639)	(32.047)	(0.214)	(0.414)
Farm size (acres)	178.495***	80.747	0.894***	0.772
	(49.970)	(61.412)	(0.327)	(0.521)
Farm size squared (acres)	-30.551	-3.501	-0.131	0.202
	(29.016)	(40.062)	(0.203)	(0.346)
Farm size cubed (acres)	1.718	-0.621	0.004	-0.027
	(4.353)	(5.004)	(0.029)	(0.054)
Farm productive assets (1,000 Ksh)	4.200**	0.826	0.026**	0.003
	(1.655)	(1.990)	(0.012)	(0.017)
Access to credit (dummy)	199.635**	138.683	1.332**	1.755
	(97.216)	(112.618)	(0.625)	(1.252)
Distance to closest market (km)	14.385**	16.603**	0.141**	0.162*
	(6.419)	(8.235)	(0.056)	(0.095)
Group official (dummy)	121.637*	101.861	0.696	1.247
	(69.198)	(108.738)	(0.623)	(0.918)
Surveyed in December (dummy)	95.791	88.435	1.314	0.796
	(111.030)	(107.989)	(0.810)	(1.198)
Poor agroecology (dummy)	151.275	112.227	1.155	-0.142
	(110.998)	(172.235)	(0.950)	(1.470)
Constant	3351.631***	800.501**	23.069***	20.418***
	(188.072)	(375.612)	(1.503)	(3.113)
Sub-county dummies	Yes	Yes	Yes	Yes
Observations	805	805	805	805
R-squared	0.295	0.045	0.192	0.128

Note: Coefficient estimates of OLS models are shown with robust standard errors clustered at farmer group level in parentheses. AE, male adult equivalent; RE, retinol equivalent. *, **, and *** significant at 10%, 5%, and 1% level, respectively. For the treatment effects, adjusted *p*-values (*q*-values) are shown in square brackets, following the two-stage procedure for multiple hypotheses testing explained by Benjamini et al. (2006).

3.6.3 Purchased and Own-Produced Foods

Tables 3.5 and 3.6 show results where the nutrition indicators are disaggregated by the consumption of calories and micronutrients from purchased and own-produced foods. The results in table 3.5 suggest that commercialization has positive and significant effects on the consumption of calories and all three micronutrients from purchased foods. A 10 percentage point increase in the level of commercialization increases calorie consumption from purchased foods by 45.9 kcal, vitamin A consumption from purchased foods by 27.4 µg, zinc consumption by 0.30 mg, and iron consumption by 0.39 mg per AE and day. An obvious interpretation is that the additional cash income generated through farm output sales improves households' economic access to food and dietary quality. Obviously, more commercialized households do not only purchase energy-dense foods, but also food items that contribute to improved micronutrient consumption, such as vegetables, fruits, and livestock products. Particularly noteworthy in table 3.5 is the positive effect of commercialization on vitamin A consumption, which is largely due to higher purchases of leafy vegetables and vitamin A-rich fruits (see table A3.11 in the appendix).

Table 3.5. Commercialization Effects on Purchased Calorie and Nutrient Consumption

Variables	Calories (kcal/day/AE)	Vitamin A (µg RE/day/AE)	Zinc (mg/day/AE)	Iron (mg/day/AE)
Commercialization (0-1)	459.274***	274.353**	2.966***	3.872***
	(163.173)	(123.787)	(1.059)	(1.404)
	[0.025]	[0.034]	[0.025]	[0.025]
Age of household head (years)	-2.104	-2.435	-0.054***	-0.059*
	(3.063)	(2.597)	(0.019)	(0.034)
Age squared (years)	0.186	0.273	0.001	0.002
	(0.193)	(0.169)	(0.001)	(0.002)
Male household head (dummy)	-31.473	-0.622	-1.340**	0.077
	(99.917)	(53.196)	(0.625)	(0.764)
Education of household head (years)	5.603	3.405	-0.001	0.149
	(11.685)	(5.741)	(0.074)	(0.107)
Household size (AE)	-209.359***	-13.164	-0.768***	-0.961***
	(23.451)	(14.243)	(0.169)	(0.275)
Farm size (acres)	-32.633	-17.760	-0.536*	-0.706
	(45.284)	(34.013)	(0.271)	(0.452)
Farm size squared (acres)	13.765	5.332	0.122	0.305
	(34.042)	(16.759)	(0.198)	(0.266)
Farm size cubed (acres)	-1.028	-0.692	-0.008	-0.033
	(4.003)	(1.777)	(0.023)	(0.030)
Farm productive assets (1,000 Ksh)	1.764	-0.063	0.002	0.011
	(1.494)	(0.809)	(0.008)	(0.010)
Access to credit (dummy)	10.233	-21.475	0.133	0.791
	(76.982)	(41.749)	(0.488)	(0.733)
Distance to closest market (km)	12.989***	2.479	0.103***	0.113
	(4.554)	(3.453)	(0.037)	(0.089)
Group official (dummy)	65.214	90.442*	0.381	0.316
	(72.339)	(51.391)	(0.501)	(0.657)
Surveyed in December (dummy)	124.183	-54.521	1.004	0.906
	(88.969)	(53.691)	(0.652)	(0.685)
Poor agroecology (dummy)	292.339**	-0.782	1.941*	1.437
	(135.146)	(64.965)	(0.975)	(1.121)
Constant	2343.893***	297.667*	13.720***	12.736***
	(215.858)	(159.097)	(1.309)	(2.569)
Sub-county dummies	Yes	Yes	Yes	Yes
Observations	805	805	805	805
R-squared	0.154	0.042	0.113	0.076

Note: Coefficient estimates of OLS models are shown with robust standard errors clustered at farmer group level in parentheses. AE, male adult equivalent; RE, retinol equivalent. *, **, and *** significant at 10%, 5%, and 1% level, respectively. For the treatment effects, adjusted *p*-values (*q*-values) are shown in square brackets, following the two-stage procedure for multiple hypotheses testing explained by Benjamini et al. (2006).

Table 3.6 shows that commercialization has no significant effects on the consumption of calories and micronutrients from own-produced foods. This is interesting, because – ceteris paribus – higher sales of farm outputs could mean lower availability of food and nutrients for home consumption. That such a decrease in the consumption of own-produced foods is not observed is likely due to higher yields on more commercialized farms. As was shown in table 3.1, the level of commercialization is positively correlated with input use and land productivity. And table A3.9 suggests that commercialization is not associated with farmers growing fewer food crops.

Table 3.6. Commercialization Effects on Own-Produced Calorie and Nutrient Consumption

Variables	Calories (kcal/day/AE)	Vitamin A (μg RE/day/AE)	Zinc (mg/day/AE)	Iron (mg/day/AE)
Commercialization (0-1)	246.271	-188.873	0.609	0.882
	(175.888)	(163.199)	(1.426)	(1.798)
	[0.118]	[0.164]	[0.337]	[0.337]
Age of household head (years)	7.820***	4.485	0.028	0.053
	(2.568)	(3.578)	(0.020)	(0.033)
Age squared (years)	0.038	-0.591***	-0.000	-0.004**
	(0.194)	(0.150)	(0.001)	(0.002)
Male household head (dummy)	-11.030	159.733**	-1.748**	0.817
	(85.307)	(71.993)	(0.732)	(0.703)
Education of household head (years)	29.904***	11.158	0.197**	0.306**
	(10.590)	(11.329)	(0.085)	(0.126)
Household size (AE)	-142.024***	-68.143***	-0.906***	-0.717**
	(25.372)	(25.121)	(0.201)	(0.291)
Farm size (acres)	226.921***	105.523***	1.462***	1.668***
	(40.637)	(33.613)	(0.303)	(0.404)
Farm size squared (acres)	-51.710*	-20.834	-0.284	-0.266
	(26.808)	(27.818)	(0.207)	(0.261)
Farm size cubed (acres)	3.489	1.581	0.015	0.024
	(3.329)	(3.702)	(0.026)	(0.039)
Farm productive assets (1,000 Ksh)	2.948**	0.581	0.027***	-0.004
	(1.400)	(2.099)	(0.010)	(0.016)
Access to credit (dummy)	181.362*	172.280*	1.272*	1.178
	(91.979)	(100.023)	(0.723)	(1.054)
Distance to closest market (km)	-0.810	11.819**	0.033	0.033
	(3.968)	(5.794)	(0.041)	(0.051)
Group official (dummy)	56.800	1.711	0.293	0.889
	(78.570)	(85.046)	(0.605)	(0.733)
Surveyed in December (dummy)	-38.223	125.919	0.314	-0.022
	(88.461)	(93.919)	(0.608)	(0.949)
Poor agroecology (dummy)	-79.972	151.220	-0.447	-1.041
	(98.164)	(148.066)	(0.834)	(1.101)
Constant	837.777***	507.902	8.484***	6.353**
	(208.657)	(307.657)	(1.475)	(2.113)
Sub-county dummies	Yes	Yes	Yes	Yes
Observations	805	805	805	805
R-squared	0.216	0.048	0.165	0.097

Note: Coefficient estimates of OLS models are shown with robust standard errors clustered at farmer group level in parentheses. AE, male adult equivalent; RE, retinol equivalent. *, **, and *** significant at 10%, 5%, and 1% level, respectively. For the treatment effects, adjusted *p*-values (*q*-values) are shown in square brackets, following the two-stage procedure for multiple hypotheses testing explained by Benjamini et al. (2006).

These results imply that commercialization does not lead to a simple substitution of purchased foods for own-produced foods. More commercialized households rather add purchased foods to their diets, without reducing the consumption of own-produced foods. This is likely due to persistent market failures. Maintaining a certain level of subsistence is a typical response of households to reduce vulnerability to market risk. Moreover, high transaction costs lead to a wide price band between selling and buying prices that may outweigh the potential benefits of stronger specialization, so it makes economic sense for households to continue producing certain commodities for home consumption (Key et al. 2000).

3.6.4 Income and Gender Pathways

The positive effects of commercialization on the consumption of calories and micronutrients from purchased foods suggest that the cash income pathway plays an important role. This is now analyzed more explicitly in table 3.7. The first column in table 3.7 reveals a significantly positive association between the level of commercialization and household income. Controlling for other factors, a 10 percentage point rise in the level of commercialization is associated with almost 25 thousand Ksh higher income (27% of mean household income of the least commercialized households). The other columns in table 3.7 confirm that gains in household income are significantly associated with higher calorie and micronutrient consumption. Only for vitamin A, the association is not statistically significant.

Table 3.7. Commercialization, Household Income, and Calorie and Nutrient Consumption

Variables	Household income	Calories (kcal/day/AE)	Vitamin A (µg RE/day/AE)	Zinc (mg/day/AE)	Iron (mg/day/AE)
Commercialization (0-1)	249.139***				
	(42.186)				
Household income (1,000 Ksh)		1.022***	0.482	0.006***	0.012***
		(0.192)	(0.290)	(0.001)	(0.003)
Off-farm income (dummy)	90.057***				
	(15.572)				
Age of household head (years)	-0.283	6.073**	4.758	-0.020	0.005
	(0.461)	(2.818)	(5.034)	(0.023)	(0.045)
Age squared (years)	0.019	0.175	-0.265	0.000	-0.002
	(0.032)	(0.210)	(0.269)	(0.002)	(0.003)
Male household head (dummy)	33.086**	-91.857	93.278	-3.401***	0.433
	(14.531)	(106.832)	(116.864)	(0.830)	(1.109)
Education of household head (years)	6.421***	34.434***	11.898	0.173**	0.434***
	(1.830)	(10.832)	(13.681)	(0.080)	(0.156)
Household size (AE)	6.718*	-371.296***	-98.945***	-1.795***	-1.917***
	(3.492)	(28.922)	(32.847)	(0.202)	(0.402)
Farm size (acres)	9.139	197.854***	78.000	0.972***	0.854
	(8.607)	(48.818)	(60.752)	(0.335)	(0.518)
Farm size squared (acres)	5.301	-45.291	-9.624	-0.222	0.069
	(5.195)	(29.174)	(41.271)	(0.220)	(0.348)
Farm size cubed (acres)	0.160	2.649	0.001	0.011	-0.019
	(0.905)	(3.860)	(4.943)	(0.028)	(0.048)
Farm productive assets (1,000 Ksh)	3.110***	1.064	-0.804	0.007	-0.036*
	(0.595)	(1.668)	(2.140)	(0.013)	(0.018)
Access to credit (dummy)	26.125**	179.365*	121.088	1.154*	1.452
	(10.345)	(100.803)	(112.444)	(0.638)	(1.251)
Poor agroecology (dummy)	16.435	86.027	74.530	0.650	-0.740
	(18.644)	(111.290)	(168.858)	(0.953)	(1.540)
Constant	-227.919***	3780.066***	1030.677**	26.198***	24.435***
	(38.844)	(204.770)	(352.596)	(1.506)	(3.170)
Sub-county dummies	Yes	Yes	Yes	Yes	Yes
Observations	805	805	805	805	805
R-squared	0.384	0.297	0.040	0.183	0.139

Note: Coefficient estimates are shown with robust standard errors clustered at farmer group level in parentheses. All models were estimated with ordinary least squares. AE, male adult equivalent; RE, retinol equivalent; Ksh, Kenyan shillings. *, **, and *** significant at 10%, 5%, and 1% level, respectively.

To evaluate possible effects of commercialization on gender roles, we look at who within the household controls the revenues from farm output sales. Most households sell different crops, for which the control of revenues can vary. Revenues from cash crops are often controlled by men, whereas for food crops the situation is more diverse (Fischer and Qaim 2012a). Hence, calculation of a single variable that captures gendered revenue control across households and crops is not straightforward. For this part of the analysis, we decided to focus on two of the most important food crops in the study region, namely maize and beans. Most of the sample households grow these crops primarily for home consumption, 25-30% of the households also sell some maize and beans to generate cash income. We only focus on the subsample of households that sold some of their maize and beans during the 12-months period covered by the survey.

For both crops, the question “who controls the revenues?” was asked with three possible and mutually exclusive answers, namely “male control”, “female control”, or “joint control”. In all cases, the answers were clear and straightforward, regardless of whether the respondent was male or female. Based on these data, we constructed separate dummy variables for both crops that take a value of one if a male household member controls the revenues alone, and zero if a female member controls the revenues either alone or jointly with a male member. This “male control” dummy takes a value of one in 23% of the households for maize, and in 17% of the households for beans.

Table 3.8 presents estimation results of models with this “male control” dummy as dependent variable for the case of maize (for beans, the models are presented in table A3.15 in the appendix). Two specifications are shown, a linear probability and a probit model. Both specifications lead to similar results. The level of commercialization is positively and significantly associated with male control of revenues in those households that sell at least some of their maize (beans). This is consistent with earlier research showing that commercialization can be associated with women losing control of how to use crop harvest and income (von Braun and Kennedy 1994; Chege et al. 2015).

Table 3.8. Association between Maize Commercialization and Male Control of Maize Revenue

Variables	Linear probability model Male controls maize revenue Coefficients	Probit model Male controls maize revenue Coefficients	 Marginal effects
Maize commercialization (0-1)	0.326***	1.312***	0.365***
	(0.112)	(0.451)	(0.122)
Age of household head (years)	-0.002	-0.009	-0.002
	(0.002)	(0.009)	(0.003)
Male household head (dummy)	0.720***	6.861***	0.684***
	(0.143)	(0.554)	(0.041)
Education of household head (years)	-0.012	-0.066	-0.018
	(0.012)	(0.046)	(0.013)
Household head married (dummy)	-0.679***	-6.587***	-0.988***
	(0.111)	(0.391)	(0.009)
Constant	0.240	-5.164***	
	(0.195)	(0.667)	
Sub-county dummies	Yes	Yes	Yes
Observations	191	191	191
R-squared	0.172		
Log pseudo likelihood		-81.729	
Pseudo R-squared		0.177	

Note: Robust standard errors clustered at farmer group level are shown in parentheses. Only maize-selling households were included. *, **, and *** significant at 10%, 5%, and 1% level, respectively.

For the same subsample of maize-selling households, table 3.9 shows that male control of revenues is associated with lower consumption of calories, vitamin A, and zinc from purchased foods (for bean-selling households, only the effect for vitamin A is statistically significant, table A3.16). In other words, women spend more on food and dietary quality than men, which seems especially relevant for vitamin A. As the models control for total household income, this negative gender pathway is a partial effect, which does not imply that the total effect of commercialization on nutrition is negative. But the analysis suggests that the total nutrition effects of commercialization could have been more positive if the loss of female control of revenues was prevented.

Table 3.9. Household Income, Gender Roles, and Consumption of Purchased Calories and Nutrients

Variables	Calories (kcal/day/AE)	Vitamin A (μg RE/day/AE)	Zinc (mg/day/AE)	Iron (mg/day/AE)
Household income (1,000 Ksh)	1.100***	0.671***	0.005***	0.008**
	(0.173)	(0.209)	(0.001)	(0.003)
Male control of maize revenue (dummy)	-314.030*	-233.409**	-1.850*	-0.271
	(160.900)	(113.886)	(1.069)	(2.026)
Age of household head (years)	1.634	-6.659	-0.064	-0.079
	(8.635)	(8.269)	(0.060)	(0.113)
Male household head (dummy)	-90.309	247.160**	-1.432	-1.255
	(162.464)	(103.838)	(1.161)	(1.779)
Education of household head (years)	-15.785	-21.426	-0.338**	-0.162
	(18.603)	(17.628)	(0.147)	(0.227)
Household size (AE)	-236.824***	-41.873	-0.911***	-1.433**
	(32.041)	(32.439)	(0.275)	(0.691)
Farm size (acres)	29.035	14.426	-0.021	0.139
	(82.937)	(45.175)	(0.578)	(0.911)
Farm size squared (acres)	24.072	0.426	0.176	-0.006
	(50.486)	(34.428)	(0.311)	(0.483)
Farm size cubed (acres)	-3.212	-0.966	-0.019	-0.006
	(5.353)	(4.466)	(0.037)	(0.054)
Farm productive assets (1,000 Ksh)	-1.526	-3.450*	-0.007	-0.003
	(2.719)	(1.795)	(0.012)	(0.023)
Access to credit (dummy)	-157.467	148.089	-1.058	1.049
	(192.238)	(88.888)	(1.133)	(2.036)
Distance to closest market (km)	18.947	-5.054	0.098	-0.085
	(18.854)	(10.045)	(0.146)	(0.160)
Group official (dummy)	180.032	29.078	0.950	0.289
	(150.986)	(110.044)	(1.083)	(1.734)
Surveyed in December (dummy)	246.329	70.259	1.746	0.905
	(229.295)	(228.718)	(1.897)	(3.036)
Poor agroecology (dummy)	296.658**	-60.207	1.954*	1.340
	(131.591)	(138.604)	(0.989)	(1.947)
Constant	2359.941***	571.188	17.336***	19.730*
	(519.677)	(551.268)	(3.705)	(8.698)
Sub-county dummies	Yes	Yes	Yes	Yes
Observations	191	191	191	191
R-squared	0.272	0.150	0.216	0.119

Note: Coefficient estimates are shown with robust standard errors clustered at farmer group level in parentheses. All models were estimated with ordinary least squares. Only maize-selling households were included. AE, male adult equivalent; RE, retinol equivalent. *, **, and *** significant at 10%, 5%, and 1% level, respectively.

3.6.5 Robustness Checks

To test the robustness of the results, we re-estimate all models in tables 3.4-3.6 using alternative indicators of commercialization. The estimated treatment effects for these alternative indicators are summarized in table 3.10 (full estimation results are shown in tables A3.17-A3.25). Regardless of the commercialization indicator used, the effects on total calorie, zinc, and iron consumption are positive and significant. The other results are also similar to the ones obtained with the original commercialization index. In most cases, higher levels of commercialization increase the consumption of calories and micronutrients from purchased foods, without significantly decreasing the consumption of own-produced foods. This underlines the robustness of the estimation results to changes in the commercialization indicator. The same conclusions are drawn when calculating the original commercialization index with ward-level prices instead of total sample mean prices to value crop and livestock outputs (table A3.26).

Beyond underlining the robustness of the findings, the estimates in table 3.10 provide a few additional insights. For maize commercialization, the positive effects on calorie, zinc, and iron consumption from own-produced foods are stronger than those from purchased foods. Maize is the main staple food, so it is not surprising that higher maize production also leads to higher consumption of this crop, especially in undernourished households. Another interesting finding is the positive and significant effect of livestock commercialization on vitamin A from own-produced foods. Meat and eggs are important sources of vitamin A, and households that produce and sell more of these products also tend to consume additional quantities. This is noteworthy, because vitamin A consumption is less responsive to income growth than the consumption of most other nutrients, and thus more difficult to influence through policies. The estimates suggest that the promotion of livestock production and marketing could be a good entry point for reducing vitamin A malnutrition.

Table 3.10. Commercialization Effects on Calorie and Nutrient Consumption with Alternative Commercialization Indicators

Variables	Total calories (kcal/day/AE)	Purchased calories (kcal/day/AE)	Own-prod. calories (kcal/day/AE)	Total vitamin A (µg RE/day/AE)	Purchased vitamin A (µg RE/day/AE)	Own-prod. vit. A (µg RE/day/AE)
Overall commercialization (0-1)	679.945***	459.274***	246.271	136.876	274.353**	-188.873
(original index)	(204.332)	(163.173)	(175.888)	(227.202)	(123.787)	(163.199)
Crop commercialization (0-1)	596.679***	358.654**	254.829*	53.892	210.803**	-200.993
	(169.927)	(157.440)	(146.849)	(202.837)	(104.165)	(154.146)
Livestock commercialization (0-1)	369.052**	308.943*	99.016	354.633*	120.867	278.481**
	(165.699)	(157.776)	(127.430)	(184.891)	(104.713)	(132.525)
Maize commercialization	997.536***	100.740	876.387***	123.636	141.793	-15.840
	(182.554)	(157.101)	(150.342)	(351.688)	(157.427)	(247.754)

Variables	Total zinc (mg/day/AE)	Purchased zinc (mg/day/AE)	Own-prod. zinc (mg/day/AE)	Total iron (mg/day/AE)	Purchased iron (mg/day/AE)	Own-prod. iron (mg/day/AE)
Overall commercialization (0-1)	3.447**	2.966***	0.609	5.546**	3.872***	0.882
(original index)	(1.307)	(1.059)	(1.426)	(2.220)	(1.404)	(1.798)
Crop commercialization (0-1)	3.145**	2.765**	0.503	4.274**	3.089**	0.438
	(1.218)	(1.084)	(1.168)	(1.873)	(1.169)	(1.571)
Livestock commercialization (0-1)	2.114*	0.910	1.386*	4.549***	3.107**	2.000
	(1.125)	(1.016)	(0.824)	(1.593)	(1.351)	(1.344)
Maize commercialization (0-1)	4.946***	0.283	4.693***	12.617***	4.168*	7.970***
	(1.647)	(1.177)	(1.168)	(2.914)	(2.252)	(2.223)

Note: Coefficient estimates are shown with robust standard errors clustered at farmer group level in parentheses. All models were estimated with ordinary least squares. Only commercialization effects are shown. Full model results are provided in tables A3.17-A3.25 in the appendix. AE, male adult equivalent; RE, retinol equivalent. *, **, and *** significant at 10%, 5%, and 1% level, respectively.

3.6.6 Continuous Treatment Effects

We now estimate continuous treatment effects with the generalized propensity score (GPS) approach. On the one hand, this helps to further test the robustness of the findings. On the other hand, accounting for the possibility of non-linear effects can also provide additional insights. Results of the GLM model with the level of commercialization as dependent variable are shown in table A3.27 in the appendix. This model is used to calculate the propensity scores. Table A3.28 shows covariate balancing tests, comparing four different treatment groups that vary in their level of commercialization. Before matching, many of the covariates for these four groups differ significantly. After matching, most of the differences turn insignificant. For the GPS analysis, we excluded 21 untreated households to avoid misleading results (Guardabascio and Ventura 2014).

Figures 3.3-3.6 present the estimated dose-response functions.[13] The consumption of total calories, zinc, and iron increases continuously with the level of commercialization, which is consistent with the parametric results discussed above. For zinc, a consumption maximum is reached at a commercialization level of about 0.7 (figure 3.5). Yet, this maximum is above the recommended minimum consumption of 15.0 mg of zinc per day, so a slight reduction beyond that point is not of nutritional concern. For calories, zinc, and iron, the consumption increases from purchased foods are also continuous, whereas the consumption from own-produced foods follows an inverse U-shape with increasing levels of commercialization. It is plausible that the benefits of subsistence, such as avoiding high transaction costs and market risk, or preferences for home-produced foods, are losing in relative importance at higher levels of commercialization and that economies of scale are starting to play a larger role.

We now turn to the discussion of the vitamin A results, which are different from those for calories, zinc, and iron. The parametric results above did not find a significant effect of commercialization on total vitamin A consumption. The non-parametric results in figure 3.4 provide interesting additional insights. The non-linear dose-response function in the left panel of figure 3.4 shows that total vitamin A consumption decreases at low levels of treatment exposure (commercialization), whereas for commercialization levels above 0.5 positive treatment effects are observed. The middle and right panels of figure 3.4 explain this non-linear effect: at low levels of treatment exposure, the decrease in vitamin A consumption from

[13] That the functions for total calorie and micronutrient consumption all lie above the respective minimum consumption thresholds, even at low levels of commercialization, should not lead to the conclusion that nutritional deficiencies are not a problem among sample households. The dose-response functions are average estimates at each level of commercialization. The descriptive statistics in table 3.2 showed that nutritional deficiencies are widespread in spite of sample mean consumption levels being above the minimum thresholds.

own-produced foods is stronger than the increase from purchased foods. This comparison is reversed at higher levels of commercialization. These results clearly suggest that vitamin A nutrition receives special attention during the process of commercialization.

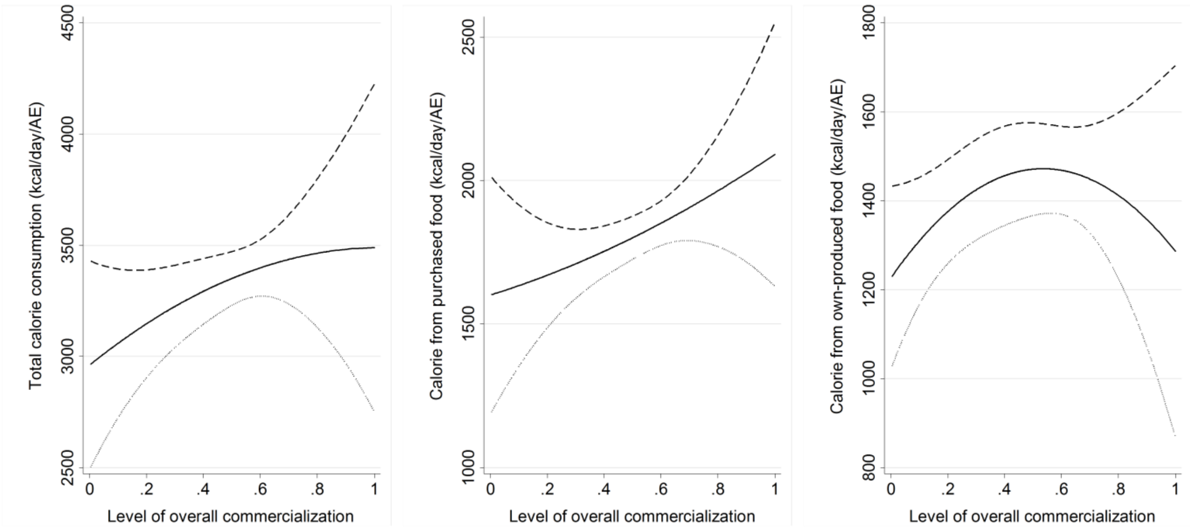

Figure 3.3. Dose-Response Functions for Commercialization Effects on Calorie Consumption

Note: Solid lines, estimated average dose-response functions; dashed lines, 95 % confidence upper bound; tight dotted lines, 95% confidence lower bound. Intervals obtained through bootstrapping with 10 replications.

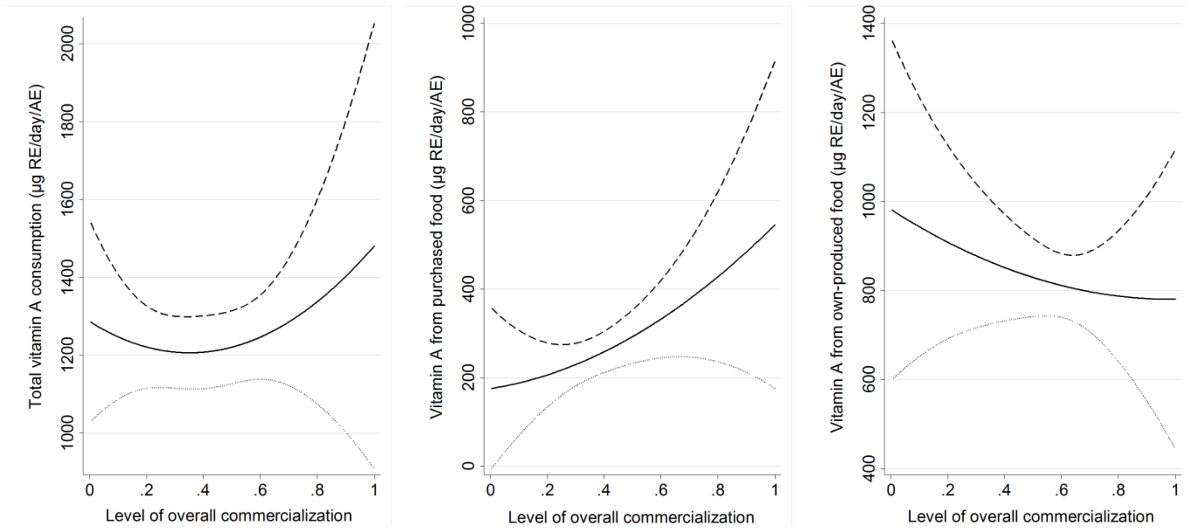

Figure 3.4. Dose-Response Functions for Commercialization Effects on Vitamin A Consumption

Note: Solid lines, estimated average dose-response functions; dashed lines, 95 % confidence upper bound; tight dotted lines, 95% confidence lower bound. Intervals obtained through bootstrapping with 10 replications.

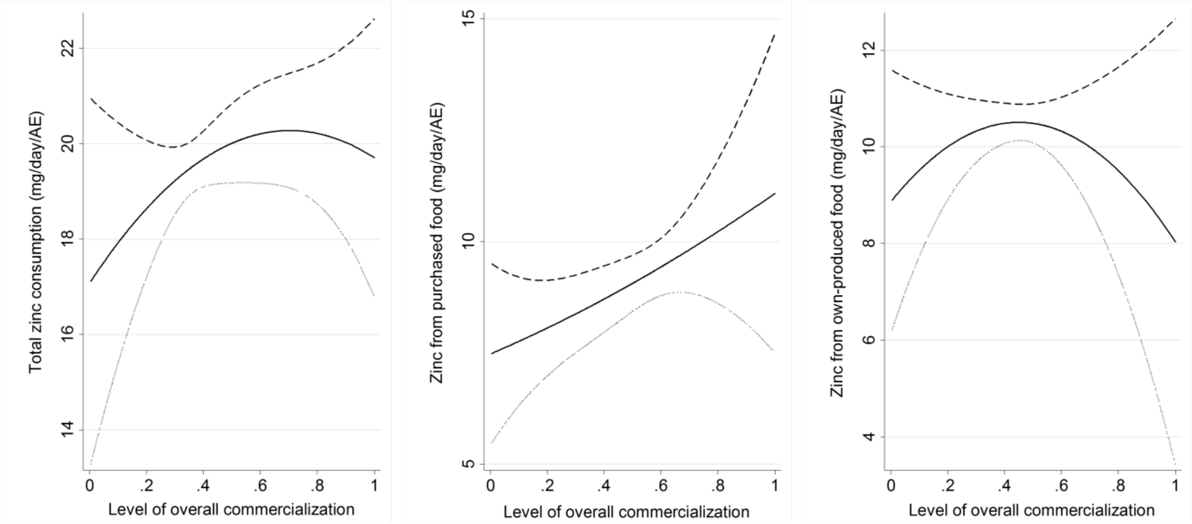

Figure 3.5. Dose-Response Functions for Commercialization Effects on Zinc Consumption

Note: Solid lines, estimated average dose-response functions; dashed lines, 95 % confidence upper bound; tight dotted lines, 95% confidence lower bound. Intervals obtained through bootstrapping with 10 replications.

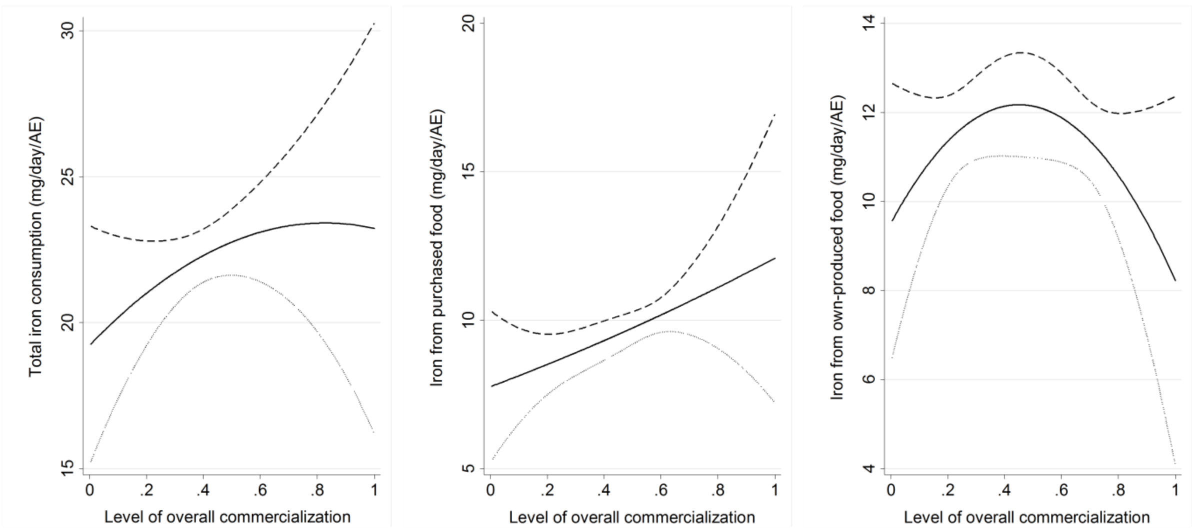

Figure 3.6. Dose-Response Functions for Commercialization Effects on Iron Consumption

Note: Solid lines, estimated average dose-response functions; dashed lines, 95 % confidence upper bound; tight dotted lines, 95% confidence lower bound. Intervals obtained through bootstrapping with 10 replications.

3.7 Conclusion

Previous studies showed that commercialization can improve productivity and income for smallholder farmers. Effects of commercialization on smallholder nutrition are less understood. Very few studies looked at this relationship, and those that did rendered mixed results (von Braun and Kennedy 1994; Carletto et al. 2017). While von Braun and Kennedy showed positive nutrition effects of commercialization, Carletto et al. (2017) mostly found insignificant effects. Bellemare and Novak (2017) found positive effects of contract farming on food security, but they differentiated between farmers with and without contracts for particular crops, not between different levels of commercialization more broadly. None of the existing studies looked explicitly at dietary quality, as we did here. Furthermore, we added to the literature by not only looking at average treatment effects, but also analyzing transmission channels.

Using survey data from smallholder farm households in Kenya, we showed that commercialization has positive effects on food security and dietary quality. Higher levels of commercialization significantly contribute to improved calorie, zinc, and iron consumption. For vitamin A consumption, the effects of commercialization were found to be insignificant. The positive effects for most dietary indicators are primarily due to rising cash incomes, allowing households to purchase more food from the market. However, rather than substituting for own-produced foods, purchased foods are added to the diet with increasing levels of commercialization. Hence, commercialization contributes to higher levels of dietary diversity. That commercialized households continue to also rely on own-produced foods is probably attributable to persistent market failures. Maintaining a certain level of subsistence is a typical response to market risks and high transaction costs. Only for highly commercialized households, the role of own-produced foods in the diets starts to decline.

We also analyzed how commercialization may affect gender roles within farm households. Commercialization increases the likelihood of male control of revenue, and this shift from female to male control has negative partial effects on the consumption of calories and micronutrients, especially vitamin A. These results confirm earlier research showing that women tend to spend more on dietary quality than men (Hoddinott and Haddad 1995; Fischer and Qaim 2012a).

Overall, we conclude that commercialization can contribute to improved nutrition in the small farm sector. An important policy implication is that enhancing market access is a key strategy to make smallholder agriculture more nutrition-sensitive. The role of women should receive particular attention. The evidence suggests that women may lose decision-making power with increasing levels of commercialization, but this may possibly be prevented through more gender-sensitive approaches and awareness-building initiatives (Meemken and Qaim 2018). We also stress that commercialization alone will not suffice to address all types of malnutrition. Commercialization helps to increase cash income, but the consumption of certain micronutrients – such as vitamin A – does not seem to be particularly responsive to income growth. Hence, more specific, complementary interventions may also be needed.

While several tests confirmed the robustness of our findings, a few limitations remain. First, the analysis relies on cross-section observational data, which limits the strength of the identification strategy. Nor do cross-section data allow the analysis of longer-term effects, which is a drawback because welfare impacts may vary over time (Carletto et al. 2010; Carletto et al. 2011). Follow-up studies with panel data and observed changes in the level of commercialization over time would be very useful. Second, the 7-day food consumption recall data provide a reasonable snapshot of dietary quality at the household level, but they do not account for seasonality and intra-household food distribution. Although we showed that seasonal variations in diets are relatively small and that the household-level nutrition indicators are significantly correlated with individual-level measures, the collection and use of higher-frequency, individual-level nutrition data would be very useful for more detailed analyses. Third, the use of 12-months recall data for farm production and marketing activities is likely associated with certain levels of imprecision. Also in this respect, higher-frequency data collected in various seasons of a year could reduce possible measurement errors. Fourth, while we tried to analyze possible effects of commercialization on gender roles within the household, a more rigorous analysis of the gender transmission channel would benefit from a larger number of gender-disaggregated variables.

A final issue that deserves discussion is that of external validity. While the concrete results are context-specific and should not simply be generalized, we argue that some broader lessons can probably be learned. The sample from Western Kenya consists of farm households with very small areas of land on which various food and cash crops are grown. Market access is limited due to poor infrastructure and inefficient institutions. Much of the food crop production is kept for home consumption, and food insecurity and nutritional deficiencies are relatively widespread. These are all characteristics that are typical for the African small farm sector, so the nutrition effects of commercialization may also not be completely different. One characteristic of the study region in Western Kenya that is more location-specific is the fact that seasonality in agricultural production and consumption is not very pronounced. This is related to ample rainfall in various months of each year. Effects of commercialization may be different in regions with stronger seasonality and higher risk of drought. The fact that our sample was drawn from households that are organized in farmer groups should be mentioned, but is unlikely to reduce external validity in a significant way. We focused on farmer groups because this allowed us to randomly sample from existing lists in the absence of county and village census data. According to own field observations, the households that are organized in farmer groups are not notably different from other farm households living in the study region.

Appendix A3

Table A3.1. Commercialization Effects on Total Calorie and Nutrient Consumption

Variables	Control function models			
	(1) Calories (kcal/day/AE)	(2) Vitamin A (µg RE/day/AE)	(3) Zinc (mg/day/AE)	(4) Iron (mg/day/AE)
Commercialization	782.610***	519.153*	5.760***	4.312
	(286.969)	(295.018)	(2.015)	(3.311)
	[0.038]	[0.107]	[0.038]	[0.153]
Age of household head (years)	5.173	3.554	-0.027	0.005
	(3.775)	(4.434)	(0.028)	(0.046)
Age squared (years)	0.343	0.004	0.002	-0.002
	(0.268)	(0.297)	(0.002)	(0.003)
Male household head (dummy)	-75.411	79.421	-3.446***	0.873
	(113.876)	(104.521)	(0.878)	(1.170)
Education of household head (years)	35.370***	5.589	0.146	0.517***
	(12.810)	(15.271)	(0.095)	(0.161)
Household size (AE)	-357.678***	-77.199**	-1.639***	-1.935***
	(33.936)	(31.902)	(0.236)	(0.419)
Farm size (acres)	150.444**	-23.702	0.262	1.110
	(76.274)	(88.876)	(0.516)	(0.943)
Farm size squared (acres)	-23.002	24.608	0.039	0.111
	(34.017)	(52.382)	(0.238)	(0.449)
Farm size cubed (acres)	1.069	-3.037	-0.011	-0.019
	(5.494)	(8.178)	(0.036)	(0.066)
Farm productive assets (1,000 Ksh)	3.992**	0.049	0.021*	0.005
	(1.917)	(2.338)	(0.013)	(0.020)
Access to credit (dummy)	184.045*	80.630	0.981	1.942*
	(101.906)	(125.388)	(0.748)	(1.111)
Distance to closest market (km)	14.848**	18.325*	0.151**	0.157*
	(6.770)	(10.426)	(0.061)	(0.095)
Group official (dummy)	111.908	65.637	0.477	1.364
	(83.927)	(114.759)	(0.650)	(0.947)
Surveyed in December (dummy)	93.786	80.970	1.269*	0.820
	(95.467)	(131.517)	(0.747)	(1.112)
Poor agroecology (dummy)	179.776	218.351	1.797*	-0.484
	(135.970)	(187.367)	(1.026)	(1.541)
Constant	3444.859***	1147.637***	25.170***	19.297***
	(357.786)	(408.825)	(2.474)	(3.907)
Sub-county dummies	Yes	Yes	Yes	Yes
Observations	805	805	805	805
R-squared	0.295	0.047	0.194	0.128
Tests of instrument validity				
p-value of motorcycles in ward = 0.018				
p-value of main market sellers in ward = 0.000				
p-value of excluded instruments (joint significance) = 0.000				
p-value of over-identification test	0.487	0.777	0.777	0.194
Bootstrap replications	500	500	500	500

Note: Coefficient estimates are shown with bootstrapped standard errors in parentheses. AE, male adult equivalent; RE, retinol equivalent. *, **, and *** significant at 10%, 5%, and 1% level, respectively. For the treatment effects, adjusted *p*-values (*q*-values) are shown in square brackets, following the two-stage procedure for multiple hypotheses testing explained by Benjamini et al. (2006).

Table A3.2. Associations between Instruments and Nutrition Indicators

Nutrition indicators	Motorcycles in ward		Main market sellers in ward	
	Correlation coefficient	Regression coefficient	Correlation coefficient	Regression coefficient
Total calories consumed (kcal/day/AE)	0.014 (0.695)	-370.226 (0.616)	0.024 (0.503)	51.725 (0.908)
Calories from purchased food (kcal/day/AE)	-0.007 (0.836)	-154.034 (0.821)	-0.000 (0.999)	86.698 (0.805)
Calories from own-produced food (kcal/day/AE)	0.024 (0.498)	-163.986 (0.812)	0.044 (0.215)	115.438 (0.744)
Total vitamin A consumed (µg RE/day/AE)	0.030 (0.403)	808.676 (0.377)	0.037 (0.298)	678.394 (0.146)
Vit. A from purchased food (µg RE/day/AE)	0.014 (0.692)	-84.923 (0.883)	0.010 (0.781)	-15.194 (0.953)
Vit. A from own-produced food (µg RE/day/AE)	0.011 (0.759)	280.770 (0.699)	0.036 (0.309)	517.035 (0.163)
Total zinc consumed (mg/day/AE)	0.056 (0.111)	2.720 (0.618)	0.044 (0.215)	3.623 (0.243)
Zinc from purchased food (mg/day/AE)	0.021 (0.551)	3.119 (0.569)	0.002 (0.964)	1.475 (0.567)
Zinc from own-produced food (mg/day/AE)	0.038 (0.283)	-2.261 (0.651)	0.046 (0.194)	1.426 (0.615)
Total iron consumed (mg/day/AE)	-0.015 (0.680)	-14.339 (0.157)	0.009 (0.792)	-4.795 (0.330)
Iron from purchased food (mg/day/AE)	0.020 (0.564)	-4.803 (0.568)	0.024 (0.501)	-0.452 (0.904)
Iron from own-produced food (mg/day/AE)	-0.040 (0.256)	-9.976 (0.188)	-0.007 (0.839)	-4.069 (0.256)

Note: The average number of motorcycles and of main market sellers in the ward are used as instruments for commercialization. *p*-values are shown in parentheses. The regression coefficients were estimated with models that include the instruments plus all other explanatory variables shown in table 3.4 of the main text.

Table A3.3. Correlation between Instruments and Mean Wealth Characteristics at Ward Level

Variables	Motorcycles in ward		Main market sellers in ward	
	Correlation coefficient	*p*-value	Correlation coefficient	*p*-value
Mean education of household head (years)	0.054	0.794	0.137	0.505
Mean household income (1,000 Ksh)	0.038	0.852	0.164	0.424
Mean farm productive assets (1,000 Ksh)	-0.039	0.851	0.083	0.687
Mean farm size (acres)	0.036	0.860	0.221	0.278

Note: The average number of motorcycles and of main market sellers in the ward are used as instruments for commercialization. Socioeconomic characteristics were computed by averaging across all sample households in the ward.

Table A3.4. Correlation between Instruments and Selected Household Socioeconomic Characteristics

Variables	Motorcycles in ward		Main market sellers in ward	
	Correlation coefficients	*p*-value	Correlation coefficients	*p*-value
Household nutrition knowledge score [a]	0.032	0.355	-0.022	0.536
Household seed expenditure per acre	-0.015	0.668	-0.047	0.184
Household fertilizer expenditure per acre	-0.025	0.477	-0.019	0.589
Household pesticide expenditure per acre	-0.057	0.106	-0.054	0.129
Household manure expenditure per acre	0.018	0.605	0.001	0.973

Note: The average number of motorcycles and of main market sellers in the ward are used as instruments for commercialization. [a] The nutrition knowledge score was computed based on four questions related to the respondents' knowledge about (i) food sources rich in vitamin A; (ii) problems of vitamin A deficiency; (iii) food sources rich in iron; and (iv) problems of iron deficiency.

Table A3.5. Mean Differences in Selected Variables between Main Market Sellers and Non-sellers

Variables	Full sample (n=805)	Main market sellers (n=258)	Main market non-sellers (n=547)	Mean difference
Value of output (1,000 Ksh)	139.382	178.152	120.990	57.162***
	(176.251)	(240.692)	(131.716)	
Value of sales (1,000 Ksh)	71.976	102.937	57.289	45.648***
	(108.139)	(146.002)	(80.552)	
Value of inputs (1,000 Ksh)	13.798	16.842	12.354	4.488***
	(14.120)	(17.214)	(12.138)	

Note: Standard deviations are shown in parentheses. Ksh, Kenyan shillings; 1 US dollar = 96.3 Ksh. *** significant at 1% level.

Table A3.6. Correlation between Household-Level and Individual-Level Nutrition Indicators

	Household level (per adult male equivalent)				
Individual level	DDS	Calories	Vitamin A	Zinc	Iron
Individual adults (n=993)					
DDS	0.275***				
Calories		0.1957***			
Vitamin A			0.147***		
Zinc				0.082**	
Iron					0.089**
Body mass index	0.049	0.103***	0.095***	0.067**	0.071**
Individual children (n=224)					
DDS	0.309***				
Calories		0.144**			
Vitamin A			0.125*		
Zinc				0.115*	
Iron					0.209***
Weight-for-age z-scores	0.132*	0.202***	-0.008	0.105	0.133*

Note: DDS, dietary diversity score with a total of 12 food groups. Pearson's correlation coefficient test was used to test for statistical significance. *, **, and *** significant at 10%, 5%, and 1% level, respectively.

Table A3.7. Correlation between Survey in December and Nutrition Indicators

Nutrition indicators	Correlation coefficient	*p*-value
Calories consumed	0.011	0.753
Vitamin A consumed	0.012	0.725
Zinc consumed	0.042	0.233
Iron consumed	0.033	0.347

Table A3.8. Correlation between Commercialization Index and other Measures of Commercialization

Variable	Correlation coefficient	*p*-value
Crop commercialization (share, 0-1)	0.900	0.000
Maize commercialization (share, 0-1)	0.313	0.000
Livestock commercialization (share, 0-1)	0.336	0.000
Value of livestock units and products sold (Ksh)	0.206	0.000
Value of crops sold (Ksh)	0.519	0.000
Value of all farm output sold (Ksh)	0.496	0.000
Value of farm output produced (Ksh)	0.281	0.000
Value of farm inputs used (Ksh)	0.310	0.000
Share of land under cash crops (0-1)	0.225	0.000

Table A3.9. Additional Summary Statistics by Level of Commercialization

Variables	Total Mean	MC25% Mean	LC25% Mean	Mean difference
Socioeconomic characteristics				
Age of household head (years)	49.27	48.35	48.34	0.01
	(12.57)	(11.22)	(13.63)	
Household size (male adult equivalents)	3.99	3.92	3.89	0.03
	(1.58)	(1.62)	(1.63)	
Number of children (5 years and below) to adults ratio	0.14	0.14	0.16	-0.02
	(0.22)	(0.24)	(0.26)	
Farm productive assets (1,000 Ksh)	19.93	23.78	15.54	8.24***
	(23.69)	(25.43)	(20.84)	
Off-farm income (dummy)	0.81	0.78	0.81	-0.04
	(0.39)	(0.42)	(0.39)	
Access to credit (dummy)	0.78	0.80	0.69	0.11**
	(0.41)	(0.40)	(0.46)	
Distance to the closest market (km)	4.91	4.60	4.97	-0.37
	(7.01)	(5.25)	(7.53)	
Distance to the closest extension agent (km)	4.34	3.89	5.52	-1.63***
	(4.93)	(4.67)	(5.40)	
Household head/spouse is a group official (dummy)	0.35	0.41	0.28	0.13***
	(0.48)	(0.49)	(0.45)	
Surveyed in December (dummy)	0.28	0.32	0.27	0.05
	(0.45)	(0.47)	(0.44)	
Poor agroecology [a] (dummy)	0.13	0.07	0.16	-0.09***
	(0.34)	(0.26)	(0.37)	
Irrigation (dummy)	0.04	0.03	0.04	-0.01
	(0.20)	(0.17)	(0.20)	
Main occupation is farming (dummy)	0.70	0.72	0.67	0.05
	(0.46)	(0.45)	(0.47)	
Motorcycles in ward [b] (number)	0.08	0.10	0.08	0.02***
	(0.05)	(0.06)	(0.05)	
Main market sellers in ward [b] (number)	0.32	0.36	0.29	0.07***
	(0.11)	(0.12)	(0.11)	
Farm production and input use				
Seed cost (1,000 Ksh/acre)	3.19	3.21	3.02	0.19
	(3.89)	(3.79)	(2.41)	
Manure cost (1,000 Ksh/acre)	0.71	0.67	0.61	0.06
	(2.96)	(2.79)	(2.17)	
Food crop production diversity (no. of food crop species)	8.01	7.99	7.56	0.42
	(3.07)	(3.30)	(3.07)	
Livestock production diversity (no. of livestock species)	3.11	3.22	2.76	0.46
	(2.97	(3.11)	(2.86)	
Farm production diversity (no. of food crop/livestock species)	11.11	11.21	10.33	0.88**
	(4.39)	(4.72)	(4.06)	
Observations	805	201	202	403

Note: Standard deviations are shown in parentheses. MC25%, 25% most commercialized households; LC25%, 25% least commercialized households; Ksh, Kenyan shillings. Value of farm productive assets excludes motorcycle to avoid possible endogeneity problems in the control function models. [a] Variable takes a value of one if farmer reported serious crop loss due to pests and diseases during last 12 months. [b] Ward-level variables were divided by number of households interviewed in each ward. *, **, and *** significant at 10%, 5%, and 1% level, respectively.

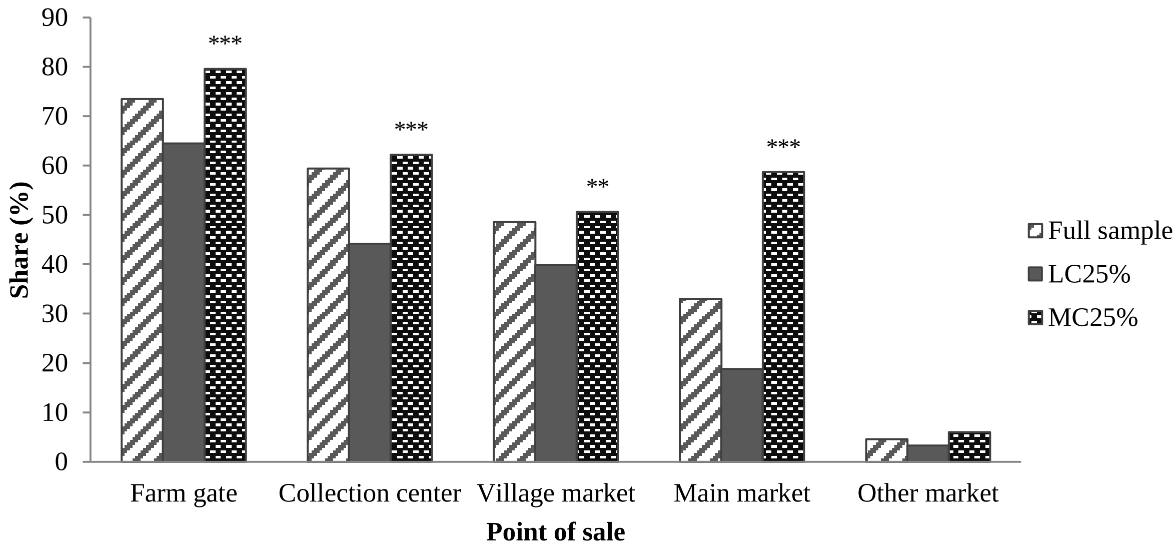

Figure A3.1. Share of Households Selling in Different Markets (Only Market Sellers Included, n=784)

Note: LC25%, 25% least commercialized households; MC25%, 25% most commercialized households. ** and *** differences between LC25% and MC25% are significant at 5% and 1% level, respectively.

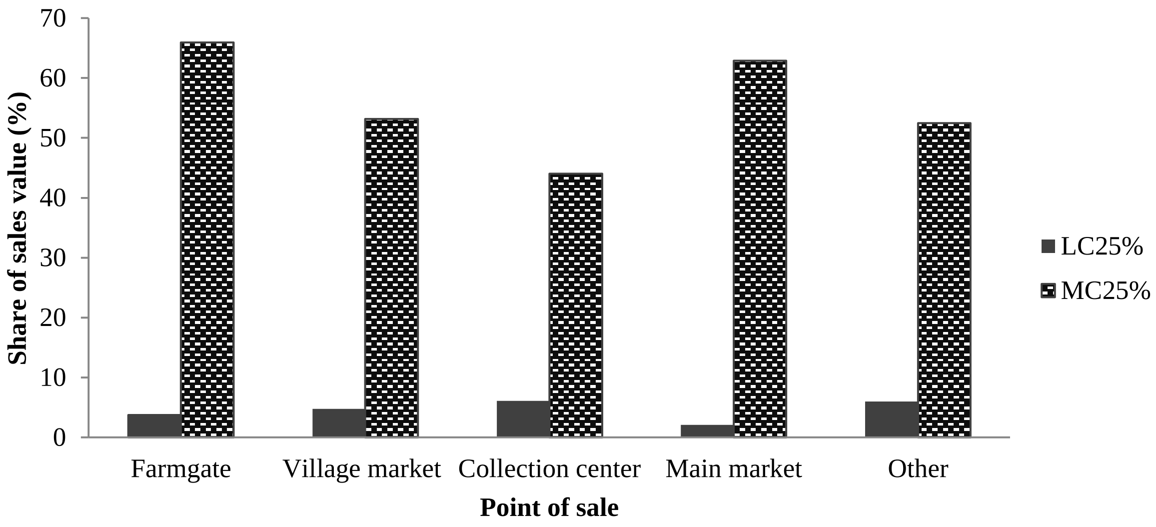

Figure A3.2. Share of Value of Sales by Level of Commercialization in Different Markets

Note: LC25%, 25% least commercialized households; MC25%, 25% most commercialized households.

Table A3.10. Overall Consumption of Micronutrients from Different Food Group by Level of Commercialization

	Total vitamin A (µg RE/day/AE)		Total zinc (mg/day/AE)		Total iron (mg/day/AE)	
Food group	LC25%	MC25%	LC25%	MC25%	LC25%	MC25%
Starchy staple foods	9.17	33.88	11.41	13.12***	8.22	11.03***
	(49.88)	(220.47)	(5.10)	(6.03)	(5.58)	(8.47)
Pulses (beans, peas, lentils)	0.60	0.75*	0.54	0.63	0.48	0.63**
	(0.81)	(0.89)	(0.68)	(0.69)	(0.67)	(0.84)
Nuts and seeds	0.00	0.01	0.04	0.04	0.01	0.01
	(0.03)	(0.11)	(0.24)	(0.22)	(0.08)	(0.06)
Dairy	125.96	146.56	1.69	1.97	0.13	0.17*
	(284.16)	(373.69)	(1.82)	(1.96)	(0.14)	(0.24)
Meat, poultry, and fish	1.27	0.76	0.82	1.18**	0.23	0.36***
	(14.64)	(5.84)	(1.68)	(1.77)	(0.46)	(0.52)
Eggs	2.24	3.33**	0.07	0.10**	0.06	0.09**
	(3.80)	(5.23)	(0.11)	(0.16)	(0.12)	(0.15)
Vitamin A-rich dark green leafy vegetables	761.47	807.77	2.02	1.96	4.32	5.68**
	(1176.40)	(1238.49)	(3.20)	(2.66)	(4.96)	(8.20)
Other vitamin A-rich fruits and vegetables	113.93	238.07**	0.15	0.20*	0.54	0.77**
	(271.12)	(773.09)	(0.32)	(0.31)	(1.03)	(1.25)
Other vegetables	82.28	103.14*	0.92	1.05	3.63	4.81**
	(118.89)	(127.83)	(1.53)	(1.28)	(4.08)	(5.45)
Other fruits	48.93	71.09**	0.60	0.79**	0.89	1.05
	(60.50)	(113.71)	(0.76)	(1.18)	(1.49)	(1.49)
Total micronutrients	1145.98	1405.35*	18.26	21.03***	18.53	24.61***
	(1231.45)	(1542.54)	(7.65)	(8.72)	(9.54)	(15.04)
Observations	202	201	202	201	202	201

Note: Mean values are shown with standard deviations in parentheses. The food group disaggregation is the one also used for the dietary diversity score with 10 food groups (DDS10) in the main paper. LC25%, 25% least commercialized households; MC25%, 25% most commercialized households; RE, retinol equivalent; AE, adult male equivalent. *, **, and *** differences between LC25% and MC25% are significant at 10%, 5%, and 1% level, respectively.

Table A3.11. Consumption of Purchased Micronutrients from Different Food Groups by Level of Commercialization

Food group	Purchased vitamin A (µg RE/day/AE) LC25%	Purchased vitamin A (µg RE/day/AE) MC25%	Purchased zinc (mg/day/AE) LC25%	Purchased zinc (mg/day/AE) MC25%	Purchased iron (mg/day/AE) LC25%	Purchased iron (mg/day/AE) MC25%
Starchy staple foods	1.40	2.29**	6.30	6.66	4.81	6.34***
	(2.94)	(4.00)	(5.59)	(5.43)	(4.97)	(6.73)
Pulses (beans, peas, lentils)	0.25	0.20	0.21	0.15	0.20	0.16
	(0.69)	(0.55)	(0.53)	(0.36)	(0.56)	(0.46)
Nuts and seeds	0.00	0.00	0.01	0.02	0.00	0.01
	(0.03)	(0.00)	(0.09)	(0.13)	(0.02)	(0.03)
Dairy	31.61	36.90	0.49	0.51	0.04	0.05
	(56.51)	(67.69)	(0.88)	(0.93)	(0.09)	(0.11)
Meat, poultry, and fish	1.28	1.67	0.71	0.99*	0.20	0.30**
	(14.68)	(14.55)	(1.54)	(1.64)	(0.41)	(0.47)
Eggs	0.74	0.78	0.02	0.02	0.02	0.02
	(2.35)	(2.61)	(0.07)	(0.07)	(0.07)	(0.09)
Vitamin A-rich dark green leafy vegetables	110.55	222.85**	0.38	0.48	1.47	2.03
	(341.06)	(717.80)	(1.15)	(1.54)	(3.61)	(6.29)
Other vitamin A-rich fruits and vegetables	54.99	94.35*	0.06	0.06	0.24	0.27
	(150.66)	(302.53)	(0.14)	(0.13)	(0.53)	(0.49)
Other vegetables	20.54	20.80	0.23	0.30	1.22	1.58*
	(67.03)	(41.72)	(0.45)	(0.52)	(1.84)	(2.43)
Other fruits	10.30	17.60	0.13	0.22	0.15	0.24
	(24.74)	(70.59)	(0.30)	(0.92)	(0.36)	(0.85)
Total micronutrients	231.65	397.31***	8.53	9.41	8.36	11.00***
	(394.53)	(788.89)	(6.85)	(6.71)	(7.65)	(10.98)
Observations	202	201	202	201	202	201

Note: Mean values are shown with standard deviations in parentheses. The food group disaggregation is the one also used for the dietary diversity score with 10 food groups (DDS10) in the main paper. LC25%, 25% least commercialized households; MC25%, 25% most commercialized households; RE, retinol equivalent; AE, adult male equivalent. *, **, and *** differences between LC25% and MC25% are significant at 10%, 5%, and 1% level, respectively.

Table A3.12. Consumption of Own-Produced Micronutrients from Different Food Groups by Level of Commercialization

	Own-produced vitamin A (µg RE/day/AE)		Own-produced zinc (mg/day/AE)		Own-produced iron (mg/day/AE)	
Food group	LC25%	MC25%	LC25%	MC25%	LC25%	MC25%
Starchy staple foods	7.75	31.44	4.95	6.35**	3.22	4.35**
	(49.99)	(220.28)	(5.25)	(5.99)	(4.29)	(6.29)
Pulses (beans, peas, lentils)	0.34	0.55***	0.31	0.47**	0.27	0.46***
	(0.54)	(0.80)	(0.53)	(0.68)	(0.44)	(0.79)
Nuts and seeds	0.00	0.01	0.02	0.02	0.01	0.01
	(0.00)	(0.11)	(0.23)	(0.19)	(0.08)	(0.05)
Dairy	87.54	108.95	1.07	1.41*	0.08	0.12
	(288.44)	(378.35)	(1.84)	(2.10)	(0.14)	(0.23)
Meat, poultry, and fish	0.00	0.00	0.08	0.19	0.03	0.06*
	(0.03)	(0.00)	(0.64)	(0.73)	(0.19)	(0.22)
Eggs	1.47	2.55***	0.04	0.08***	0.04	0.07**
	(3.29)	(4.89)	(0.10)	(0.15)	(0.10)	(0.13)
Vitamin A-rich dark green leafy vegetables	554.37	528.70	1.48	1.37	2.55	3.30*
	(943.04)	(796.66)	(2.62)	(1.97)	(3.40)	(5.18)
Other vitamin A-rich fruits and vegetables	51.36	104.81	0.08	0.12*	0.27	0.45**
	(181.40)	(493.06)	(0.25)	(0.25)	(0.75)	(1.02)
Other vegetables	53.02	81.63***	0.64	0.69	2.13	3.04**
	(92.12)	(125.95)	(1.50)	(1.05)	(3.61)	(4.59)
Other fruits	31.00	42.97*	0.37	0.49	0.61	0.66
	(51.87)	(84.37)	(0.61)	(0.86)	(1.46)	(1.22)
Total micronutrients	786.85	901.60	9.06	11.19***	9.20	12.52***
	(1000.91)	(1136.97)	(6.86)	(8.04)	(7.17)	(11.44)
Observations	202	201	202	201	202	201

Note: Mean values are shown with standard deviations in parentheses. The food group disaggregation is the one also used for the dietary diversity score with 10 food groups (DDS10) in the main paper. LC25%, 25% least commercialized households; MC25%, 25% most commercialized households; RE, retinol equivalent; AE, adult male equivalent. *, **, and *** differences between LC25% and MC25% are significant at 10%, 5%, and 1% level, respectively.

Table A3.13. Commercialization Effects on Purchased Calorie and Nutrient Consumption

	Control function models			
Variables	(1) Calories (kcal/day/AE)	(2) Vitamin A (µg RE/day/AE)	(3) Zinc (mg/day/AE)	(4) Iron (mg/day/AE)
Commercialization	551.941**	277.164**	3.565**	4.344*
	(257.880)	(135.539)	(1.666)	(2.366)
	[0.087]	[0.090]	[0.087]	[0.107]
Age of household head (years)	-2.428	-2.445	-0.056**	-0.061*
	(3.478)	(2.109)	(0.026)	(0.031)
Age squared (years)	0.244	0.274	0.001	0.002
	(0.244)	(0.171)	(0.002)	(0.002)
Male household head (dummy)	-38.579	-0.838	-1.386**	0.041
	(96.775)	(50.791)	(0.651)	(0.745)
Education of household head (years)	3.505	3.341	-0.014	0.139
	(11.291)	(6.546)	(0.079)	(0.110)
Household size (AE)	-203.607***	-12.989	-0.731***	-0.932***
	(27.290)	(15.299)	(0.165)	(0.251)
Farm size (acres)	-57.953	-18.528	-0.700	-0.835
	(68.454)	(40.768)	(0.455)	(0.717)
Farm size squared (acres)	20.579	5.539	0.167	0.339
	(32.210)	(17.939)	(0.208)	(0.313)
Farm size cubed (acres)	-1.613	-0.710	-0.012	-0.036
	(4.672)	(2.024)	(0.030)	(0.038)
Farm productive assets (1,000 Ksh)	1.576	-0.068	0.000	0.010
	(1.485)	(0.919)	(0.009)	(0.012)
Access to credit (dummy)	-3.839	-21.902	0.042	0.719
	(90.146)	(56.098)	(0.684)	(0.795)
Distance to closest market (km)	13.407**	2.491	0.105**	0.115
	(5.998)	(3.796)	(0.043)	(0.092)
Group official (dummy)	56.433	90.176*	0.325	0.271
	(78.514)	(49.601)	(0.582)	(0.697)
Surveyed in December (dummy)	122.374	-54.576	0.992*	0.897
	(84.753)	(49.404)	(0.560)	(0.758)
Poor agroecology (dummy)	318.065**	-0.002	2.108**	1.568
	(126.197)	(73.081)	(0.918)	(1.126)
Constant	2428.042***	300.219*	14.265***	13.164***
	(324.358)	(173.742)	(2.393)	(2.984)
Sub-county dummies	Yes	Yes	Yes	Yes
Observations	805	805	805	805
R-squared	0.154	0.042	0.113	0.076
Tests of instrument validity				
p-value of motorcycles in ward = 0.018				
p-value of main market sellers in ward = 0.000				
p-value of excluded instruments (joint significance) = 0.000				
p-value of over-identification test	0.688	0.777	0.688	0.394
Bootstrap replications	500	500	500	500

Note: Coefficient estimates are shown with bootstrapped standard errors in parentheses. AE, male adult equivalent; RE, retinol equivalent. *, **, and *** significant at 10%, 5%, and 1% level, respectively. For the treatment effects, adjusted *p*-values (*q*-values) are shown in square brackets, following the two-stage procedure for multiple hypotheses testing explained by Benjamini et al. (2006).

Table A3.14. Commercialization Effects on Own-Produced Calorie and Nutrient Consumption

	Control function models			
Variables	(1) Calories (kcal/day/AE)	(2) Vitamin A (μg RE/day/AE)	(3) Zinc (mg/day/AE)	(4) Iron (mg/day/AE)
Commercialization	362.660	159.599	2.087	-0.539
	(238.082)	(229.432)	(1.763)	(2.348)
	[0.127]	[0.238]	[0.155]	[0.347]
Age of household head (years)	7.414**	3.267	0.023	0.058*
	(2.937)	(3.329)	(0.022)	(0.032)
Age squared (years)	0.111	-0.374*	0.001	-0.005**
	(0.239)	(0.200)	(0.002)	(0.002)
Male household head (dummy)	-19.954	133.014*	-1.861**	0.926
	(88.502)	(75.039)	(0.786)	(0.783)
Education of household head (years)	27.269**	3.269	0.163**	0.338***
	(11.145)	(12.416)	(0.082)	(0.114)
Household size (AE)	-134.799***	-46.512*	-0.814***	-0.805**
	(27.452)	(25.840)	(0.204)	(0.321)
Farm size (acres)	195.120***	10.310	1.058**	2.056***
	(68.106)	(66.961)	(0.487)	(0.644)
Farm size squared (acres)	-43.152	4.789	-0.176	-0.371
	(28.711)	(34.184)	(0.229)	(0.337)
Farm size cubed (acres)	2.753	-0.621	0.005	0.033
	(4.050)	(5.277)	(0.034)	(0.055)
Farm productive assets (1,000 Ksh)	2.712*	-0.126	0.024**	-0.001
	(1.618)	(1.933)	(0.011)	(0.015)
Access to credit (dummy)	163.687*	119.361	1.048	1.394
	(83.535)	(85.844)	(0.646)	(0.861)
Distance to closest market (km)	-0.286	13.389*	0.040	0.027
	(4.446)	(7.100)	(0.047)	(0.059)
Group official (dummy)	45.772	-31.310	0.153	1.024
	(71.027)	(88.485)	(0.577)	(0.747)
Surveyed in December (dummy)	-40.496	119.114	0.285	0.005
	(81.755)	(109.962)	(0.632)	(0.835)
Poor agroecology (dummy)	-47.661	247.959	-0.037	-1.435
	(106.103)	(164.788)	(0.751)	(1.112)
Constant	943.466***	824.341**	9.826***	5.063*
	(307.979)	(339.059)	(2.173)	(2.805)
Sub-county dummies	Yes	Yes	Yes	Yes
Observations	805	805	805	805
R-squared	0.217	0.051	0.166	0.097
Tests of instrument validity				
p-value of motorcycles in ward = 0.018				
p-value of main market sellers in ward = 0.000				
p-value of excluded instruments (joint significance) = 0.000				
p-value of over-identification test	0.570	0.688	0.347	0.288
Bootstrap replications	500	500	500	500

Note: Coefficient estimates are shown with bootstrapped standard errors in parentheses. AE, male adult equivalent; RE, retinol equivalent. *, **, and *** significant at 10%, 5%, and 1% level, respectively. For the treatment effects, adjusted *p*-values (*q*-values) are shown in square brackets, following the two-stage procedure for multiple hypotheses testing explained by Benjamini et al. (2006).

Table A3.15. Association between Bean Commercialization and Male Control of Bean Revenue

Variables	Linear probability model Male controls bean revenue Coefficients	Probit model Male controls bean revenue Coefficients	 Marginal effects
Bean commercialization (0-1)	0.284***	1.191**	0.271***
	(0.101)	(0.469)	(0.100)
Age of household head (years)	0.001	0.007	0.001
	(0.002)	(0.008)	(0.002)
Male household head (dummy)	0.405***	1.703***	0.219***
	(0.126)	(0.698)	(0.050)
Education of household head (years)	-0.000	-0.003	-0.001
	(0.008)	(0.035)	(0.008)
Household head married (dummy)	-0.542***	-2.134***	-0.676***
	(0.119)	(0.671)	(0.187)
Constant	0.087	-1.559**	
	(0.136)	(0.590)	
Sub-county dummies	Yes	Yes	Yes
Observations	275	275	275
R-squared	0.143		
Log likelihood		-108.437	
Pseudo R-squared		0.148	

Note: Robust standard errors clustered at farmer group level are shown in parentheses. Only bean-selling households were included. *, **, and *** significant at 10%, 5%, and 1% level, respectively.

Table A3.16. Household Income, Gender Roles, and Consumption of Purchased Calories and Nutrients

Variables	Calories (kcal/day/AE)	Vitamin A (μg RE/day/AE)	Zinc (mg/day/AE)	Iron (mg/day/AE)
Household income (1,000 Ksh)	0.516*	0.470**	0.002	0.008**
	(0.294)	(0.182)	(0.002)	(0.003)
Male control of bean revenue (dummy)	-86.712	-113.559*	-0.820	1.086
	(145.604)	(60.946)	(0.980)	(2.024)
Age of household head (years)	6.492	-3.322	-0.010	-0.056
	(6.471)	(5.458)	(0.058)	(0.080)
Male household head (dummy)	311.477**	113.510	1.372	1.909
	(142.274)	(78.466)	(1.006)	(1.579)
Education of household head (years)	-35.866**	-21.936	-0.300**	-0.203
	(17.154)	(13.766)	(0.126)	(0.211)
Household size (AE)	-264.778***	-34.329	-1.118**	-1.377**
	(44.326)	(27.995)	(0.391)	(0.564)
Farm size (acres)	28.778	69.992	0.136	0.481
	(77.684)	(62.039)	(0.501)	(0.923)
Farm size squared (acres)	-42.446	-25.141	-0.270	-0.322
	(48.463)	(35.831)	(0.303)	(0.488)
Farm size cubed (acres)	4.738	1.438	0.028	0.017
	(5.152)	(4.014)	(0.033)	(0.052)
Farm productive assets (1,000 Ksh)	1.985	-2.535	0.006	-0.015
	(2.249)	(1.514)	(0.013)	(0.022)
Access to credit (dummy)	-71.199	35.057	-0.712	0.411
	(151.674)	(71.622)	(0.891)	(1.564)
Distance to closest market (km)	18.143**	8.278	0.153*	0.310
	(7.906)	(7.120)	(0.077)	(0.228)
Group official (dummy)	91.616	55.842	0.382	-0.341
	(119.873)	(115.461)	(0.776)	(1.138)
Surveyed in December (dummy)	307.215**	-29.892	2.437**	1.157
	(142.364)	(108.401)	(1.033)	(1.461)
Poor agroecology (dummy)	334.682	7.115	2.647*	1.712
	(192.152)	(111.084)	(1.270)	(1.566)
Constant	2191.667***	446.717	12.958***	15.508**
	(367.062)	(415.876)	(3.804)	(6.744)
Sub-county dummies	Yes	Yes	Yes	Yes
Observations	275	275	275	275
R-squared	0.232	0.098	0.185	0.136

Note: Coefficient estimates are shown with robust standard errors clustered at farmer group level in parentheses. All models were estimated with ordinary least squares. Only bean-selling households were included. AE, male adult equivalent; RE, retinol equivalent. *, **, and *** significant at 10%, 5%, and 1% level, respectively.

Table A3.17. Crop Commercialization Effects on Total Calorie and Nutrient Consumption

Variables	Calories (kcal/day/AE)	Vitamin A (µg RE/day/AE)	Zinc (mg/day/AE)	Iron (mg/day/AE)
Crop commercialization (0-1)	596.679***	53.892	3.145**	4.274**
	(169.927)	(202.837)	(1.218)	(1.873)
Age of household head (years)	5.601*	4.954	-0.019	0.001
	(2.881)	(5.441)	(0.024)	(0.049)
Age squared (years)	0.278	-0.245	0.001	-0.001
	(0.217)	(0.281)	(0.002)	(0.003)
Male household head (dummy)	-59.646	110.632	-3.229***	0.845
	(103.426)	(110.110)	(0.766)	(1.113)
Education of household head (years)	38.316***	14.664	0.201**	0.497***
	(11.463)	(14.868)	(0.087)	(0.163)
Household size (AE)	-365.677***	-102.058***	-1.790***	-1.879***
	(30.945)	(32.026)	(0.213)	(0.407)
Farm size (acres)	176.877***	84.978	0.877***	0.800
	(48.269)	(62.497)	(0.317)	(0.509)
Farm size squared (acres)	-29.875	-4.585	-0.126	0.196
	(29.095)	(39.936)	(0.198)	(0.346)
Farm size cubed (acres)	1.773	-0.519	0.004	-0.025
	(4.342)	(4.999)	(0.029)	(0.054)
Farm productive assets (1,000 Ksh)	4.503**	0.886	0.028**	0.005
	(1.687)	(1.996)	(0.012)	(0.017)
Access to credit (dummy)	199.393**	140.986	1.327**	1.774
	(97.305)	(112.418)	(0.627)	(1.234)
Distance to closest market (km)	14.899**	16.575**	0.144**	0.165*
	(6.511)	(8.208)	(0.057)	(0.096)
Group official (dummy)	120.277	103.246	0.686	1.251
	(68.500)	(108.827)	(0.619)	(0.912)
Surveyed in December (dummy)	102.738	89.389	1.350	0.848
	(111.331)	(108.924)	(0.815)	(1.207)
Poor agroecology (dummy)	150.411	107.775	1.158	-0.187
	(110.662)	(173.190)	(0.945)	(1.469)
Constant	3371.385***	826.681**	23.129***	20.778***
	(195.623)	(376.683)	(1.553)	(3.145)
Sub-county dummies	Yes	Yes	Yes	Yes
Observations	805	805	805	805
R-squared	0.294	0.045	0.192	0.126

Note: Coefficient estimates are shown with robust standard errors clustered at farmer group level in parentheses. All models were estimated with ordinary least squares. AE, male adult equivalent; RE, retinol equivalent. *, **, and *** significant at 10%, 5%, and 1% level, respectively.

Table A3.18. Crop Commercialization Effects on Purchased Calorie and Nutrient Consumption

Variables	Calories (kcal/day/AE)	Vitamin A (µg RE/day/AE)	Zinc (mg/day/AE)	Iron (mg/day/AE)
Crop commercialization (0-1)	358.654**	210.803**	2.765**	3.089**
	(157.440)	(104.165)	(1.084)	(1.169)
Age of household head (years)	-2.024	-2.385	-0.054***	-0.059
	(3.067)	(2.601)	(0.020)	(0.035)
Age squared (years)	0.179	0.268	0.001	0.002
	(0.190)	(0.170)	(0.001)	(0.002)
Male household head (dummy)	-25.933	2.704	-1.307**	0.124
	(99.683)	(53.183)	(0.621)	(0.771)
Education of household head (years)	6.222	3.790	0.001	0.154
	(11.804)	(5.822)	(0.075)	(0.109)
Household size (AE)	-210.995***	-14.182	-0.774***	-0.974***
	(22.799)	(14.149)	(0.167)	(0.272)
Farm size (acres)	-30.672	-16.352	-0.554**	-0.694
	(44.795)	(33.873)	(0.265)	(0.437)
Farm size squared (acres)	13.403	5.053	0.128	0.303
	(34.552)	(16.741)	(0.198)	(0.266)
Farm size cubed (acres)	-0.929	-0.629	-0.008	-0.032
	(4.047)	(1.785)	(0.023)	(0.031)
Farm productive assets (1,000 Ksh)	1.968	0.059	0.003	0.013
	(1.512)	(0.801)	(0.008)	(0.010)
Access to credit (dummy)	11.646	-20.510	0.127	0.800
	(76.054)	(40.933)	(0.485)	(0.725)
Distance to closest market (km)	13.249***	2.627	0.105***	0.115
	(4.633)	(3.450)	(0.037)	(0.089)
Group official (dummy)	65.407	90.643*	0.371	0.316
	(71.313)	(51.530)	(0.495)	(0.657)
Surveyed in December (dummy)	128.578	-51.919	1.036	0.944
	(89.079)	(53.957)	(0.652)	(0.690)
Poor agroecology (dummy)	288.888**	-3.066	1.948*	1.412
	(136.595)	(64.288)	(0.985)	(1.111)
Constant	2372.115***	315.680*	13.752***	12.952***
	(215.565)	(160.028)	(1.310)	(2.582)
Sub-county dummies	Yes	Yes	Yes	Yes
Observation	805	805	805	805
R-squared	0.152	0.040	0.113	0.074

Note: Coefficient estimates are shown with robust standard errors clustered at farmer group level in parentheses. All models were estimated with ordinary least squares. AE, male adult equivalent; RE, retinol equivalent. *, **, and *** significant at 10%, 5%, and 1% level, respectively.

Table A3.19. Crop Commercialization Effects on Own-Produced Calorie and Nutrient Consumption

Variables	Calories (kcal/day/AE)	Vitamin A (µg RE/day/AE)	Zinc (mg/day/AE)	Iron (mg/day/AE)
Crop commercialization (0-1)	254.829*	-200.993	0.503	0.438
	(146.849)	(154.146)	(1.168)	(1.571)
Age of household head (years)	7.816***	4.492	0.028	0.053
	(2.540)	(3.567)	(0.020)	(0.033)
Age squared (years)	0.043	-0.596***	-0.000	-0.004**
	(0.194)	(0.148)	(0.001)	(0.002)
Male household head (dummy)	-8.353	157.707**	-1.740**	0.829
	(85.259)	(72.349)	(0.730)	(0.699)
Education of household head (years)	29.956**	11.143	0.198**	0.308**
	(10.569)	(11.332)	(0.084)	(0.128)
Household size (AE)	-142.144***	-68.118***	-0.908***	-0.723**
	(25.517)	(25.295)	(0.203)	(0.293)
Farm size (acres)	223.670***	108.398***	1.462***	1.689***
	(41.144)	(34.824)	(0.304)	(0.399)
Farm size squared (acres)	-50.752*	-21.672	-0.284	-0.272
	(26.683)	(27.893)	(0.207)	(0.260)
Farm size cubed (acres)	3.456	1.613	0.015	0.025
	(3.294)	(3.741)	(0.026)	(0.039)
Farm productive assets (1,000 Ksh)	3.058**	0.497	0.027***	-0.004
	(1.415)	(2.116)	(0.010)	(0.015)
Access to credit (dummy)	179.899*	173.599*	1.273*	1.190
	(92.177)	(100.032)	(0.724)	(1.054)
Distance to closest market (km)	-0.548	11.607**	0.033	0.033
	(3.970)	(5.735)	(0.041)	(0.051)
Group official (dummy)	55.339	2.971	0.293	0.896
	(78.849)	(85.266)	(0.607)	(0.732)
Surveyed in December (dummy)	-35.447	123.753	0.320	-0.016
	(89.127)	(94.604)	(0.612)	(0.951)
Poor agroecology (dummy)	-77.784	149.183	-0.450	-1.063
	(98.046)	(149.331)	(0.827)	(1.110)
Constant	831.950***	514.234	8.512***	6.491***
	(211.351)	(308.940)	(1.494)	(2.102)
Sub-county dummies	Yes	Yes	Yes	Yes
Observation	805	805	805	805
R-squared	0.217	0.049	0.165	0.096

Note: Coefficient estimates are shown with robust standard errors clustered at farmer group level in parentheses. All models were estimated with ordinary least squares. AE, male adult equivalent; RE, retinol equivalent. *, **, and *** significant at 10%, 5%, and 1% level, respectively.

Table A3.20. Maize Commercialization Effects on Total Calorie and Nutrient Consumption

Variables	Calories (kcal/day/AE)	Vitamin A (µg RE/day/AE)	Zinc (mg/day/AE)	Iron (mg/day/AE)
Maize commercialization (0-1)	997.536***	123.636	4.946***	12.617***
	(182.554)	(351.688)	(1.647)	(2.914)
Age of household head (years)	5.576**	4.544	-0.025	-0.008
	(2.474)	(5.617)	(0.023)	(0.047)
Age squared (years)	0.275	-0.223	0.001	-0.001
	(0.215)	(0.300)	(0.002)	(0.003)
Male household head (dummy)	-58.749	113.117	-3.173***	1.208
	(114.177)	(118.313)	(0.871)	(1.055)
Education of household head (years)	40.597***	12.997	0.213**	0.464***
	(11.211)	(16.112)	(0.093)	(0.160)
Household size (AE)	-351.993***	-96.498***	-1.741***	-1.693***
	(30.854)	(33.822)	(0.225)	(0.411)
Farm size (acres)	188.337***	76.766	0.952***	0.712
	(48.451)	(59.990)	(0.334)	(0.475)
Farm size squared (acres)	-40.926	1.639	-0.186	0.209
	(29.943)	(45.852)	(0.205)	(0.307)
Farm size cubed (acres)	2.646	-1.291	0.008	-0.030
	(3.881)	(5.620)	(0.027)	(0.046)
Farm productive assets (1,000 Ksh)	3.728**	1.024	0.024*	-0.003
	(1.566)	(2.168)	(0.012)	(0.016)
Access to credit (dummy)	224.364**	132.060	1.393**	2.098*
	(98.564)	(124.080)	(0.622)	(1.235)
Distance to closest market (km)	16.833**	20.100*	0.167***	0.212*
	(7.142)	(10.638)	(0.059)	(0.110)
Distance to extension agent (km)	-4.564	-10.974	-0.078*	-0.029
	(6.488)	(11.889)	(0.046)	(0.103)
Group official (dummy)	94.284	97.014	0.503	1.007
	(67.868)	(116.585)	(0.646)	(0.893)
Surveyed in December (dummy)	68.967	68.663	1.225	0.698
	(105.282)	(121.842)	(0.831)	(1.188)
Poor agroecology (dummy)	135.197	138.728	1.216	0.200
	(108.288)	(179.550)	(0.926)	(1.485)
Constant	3468.252***	887.131**	24.144***	20.910***
	(202.491)	(392.811)	(1.611)	(3.286)
Sub-county dummies	Yes	Yes	Yes	Yes
Observations	755	755	755	755
R-squared	0.312	0.044	0.207	0.172

Note: Coefficient estimates are shown with robust standard errors clustered at farmer group level in parentheses. All models were estimated with ordinary least squares. AE, male adult equivalent; RE, retinol equivalent. *, **, and *** significant at 10%, 5%, and 1% level, respectively.

Table A3.21. Maize Commercialization Effects on Purchased Calorie and Nutrient Consumption

Variables	Calories (kcal/day/AE)	Vitamin A (µg RE/day/AE)	Zinc (mg/day/AE)	Iron (mg/day/AE)
Maize commercialization (0-1)	100.739	141.793	0.283	4.168*
	(157.101)	(157.427)	(1.177)	(2.252)
Age of household head (years)	-2.135	-2.125	-0.057***	-0.065*
	(2.948)	(2.666)	(0.020)	(0.034)
Age squared (years)	0.101	0.260	0.000	0.002
	(0.183)	(0.175)	(0.001)	(0.002)
Male household head (dummy)	-31.704	5.117	-1.233*	0.107
	(106.311)	(59.809)	(0.659)	(0.774)
Education of household head (years)	13.652	3.459	0.040	0.150
	(11.277)	(6.251)	(0.075)	(0.111)
Household size (AE)	-211.830***	-13.312	-0.817***	-0.902***
	(24.792)	(15.684)	(0.178)	(0.286)
Farm size (acres)	4.951	-0.765	-0.239	-0.477
	(43.189)	(33.160)	(0.278)	(0.387)
Farm size squared (acres)	-1.231	-0.744	0.019	0.236
	(35.258)	(17.120)	(0.209)	(0.259)
Farm size cubed (acres)	0.333	-0.208	0.001	-0.029
	(4.117)	(1.831)	(0.024)	(0.030)
Farm productive assets (1,000 Ksh)	1.570	-0.191	-0.000	0.010
	(1.567)	(0.886)	(0.009)	(0.010)
Access to credit (dummy)	34.349	-26.588	0.226	0.897
	(82.678)	(44.563)	(0.529)	(0.773)
Distance to closest market (km)	14.360***	3.173	0.113***	0.136
	(4.953)	(3.885)	(0.040)	(0.101)
Distance to extension agent (km)	-9.273	-0.074	-0.059	-0.050
	(6.452)	(4.430)	(0.049)	(0.078)
Group official (dummy)	17.614	80.739	0.086	-0.044
	(68.194)	(50.605)	(0.467)	(0.593)
Surveyed in December (dummy)	85.766	-75.725	0.797	0.618
	(81.985)	(53.368)	(0.639)	(0.670)
Poor agroecology (dummy)	292.547**	-3.071	2.000**	1.569
	(136.451)	(65.396)	(0.987)	(1.159)
Constant	2472.188***	362.630**	14.741***	13.706***
	(219.935)	(175.805)	(1.388)	(2.616)
Sub-county dummies	Yes	Yes	Yes	Yes
Observations	755	755	755	755
R-squared	0.149	0.039	0.108	0.083

Note: Coefficient estimates are shown with robust standard errors clustered at farmer group level in parentheses. All models were estimated with ordinary least squares. AE, male adult equivalent; RE, retinol equivalent. *, **, and *** significant at 10%, 5%, and 1% level, respectively.

Table A3.22. Maize Commercialization Effects on Own-Produced Calorie and Nutrient Consumption

Variables	Calories (kcal/day/AE)	Vitamin A (µg RE/day/AE)	Zinc (mg/day/AE)	Iron (mg/day/AE)
Maize commercialization (0-1)	876.387***	-15.840	4.693***	7.970***
	(150.342)	(247.754)	(1.168)	(2.223)
Age of household head (years)	7.795***	3.886	0.024	0.049
	(2.619)	(3.678)	(0.019)	(0.034)
Age squared (years)	0.098	-0.569***	0.000	-0.004
	(0.204)	(0.164)	(0.001)	(0.002)
Male household head (dummy)	1.549	156.714**	-1.728**	1.259*
	(87.191)	(74.126)	(0.778)	(0.695)
Education of household head (years)	24.423**	10.487	0.165**	0.263**
	(10.016)	(12.673)	(0.082)	(0.117)
Household size (AE)	-127.388***	-63.330**	-0.811***	-0.602**
	(26.017)	(26.587)	(0.210)	(0.292)
Farm size (acres)	201.793***	84.032**	1.216***	1.277***
	(46.491)	(35.000)	(0.363)	(0.443)
Farm size squared (acres)	-46.671	-9.883	-0.226	-0.139
	(28.164)	(32.254)	(0.222)	(0.247)
Farm size cubed (acres)	3.013	0.493	0.009	0.012
	(3.168)	(4.163)	(0.026)	(0.034)
Farm productive assets (1,000 Ksh)	2.700**	0.887	0.027***	-0.008
	(1.307)	(2.226)	(0.010)	(0.015)
Access to credit (dummy)	186.075**	169.411	1.256*	1.365
	(91.956)	(109.404)	(0.741)	(1.019)
Distance to closest market (km)	-0.651	13.332*	0.046	0.047
	(4.694)	(7.484)	(0.046)	(0.060)
Distance to extension agent (km)	6.368	-6.602	-0.019	0.036
	(8.977)	(8.473)	(0.052)	(0.094)
Group official (dummy)	73.303	1.125	0.366	0.941
	(75.090)	(92.388)	(0.624)	(0.737)
Surveyed in December (dummy)	-26.898	127.992	0.429	0.142
	(95.796)	(103.926)	(0.664)	(1.015)
Poor agroecology (dummy)	-99.188	181.664	-0.457	-0.879
	(103.005)	(154.832)	(0.854)	(1.099)
Constant	837.175***	490.211	8.660***	5.847**
	(228.410)	(314.917)	(1.565)	(2.224)
Sub-county dummies	Yes	Yes	Yes	Yes
Observation	755	755	755	755
R-squared	0.236	0.046	0.174	0.126

Note: Coefficient estimates are shown with robust standard errors clustered at farmer group level in parentheses. All models were estimated with ordinary least squares. AE, male adult equivalent; RE, retinol equivalent. *, **, and *** significant at 10%, 5%, and 1% level, respectively.

Table A3.23. Livestock Commercialization Effects on Total Calorie and Nutrient Consumption

Variables	Calories (kcal/day/AE)	Vitamin A (µg RE/day/AE)	Zinc (mg/day/AE)	Iron (mg/day/AE)
Livestock commercialization (0-1)	369.052**	354.633*	2.114*	4.549***
	(165.699)	(184.891)	(1.125)	(1.593)
Age of household head (years)	4.629	3.853	-0.017	0.009
	(3.078)	(6.036)	(0.026)	(0.050)
Age squared (years)	0.253	-0.270	0.001	-0.002
	(0.248)	(0.309)	(0.002)	(0.003)
Male household head (dummy)	-113.786	187.927*	-3.059***	0.857
	(114.909)	(109.640)	(0.963)	(1.115)
Education of household head (years)	32.770***	8.749	0.159*	0.381**
	(11.912)	(15.294)	(0.090)	(0.158)
Household size (AE)	-370.274***	-102.261***	-1.867***	-1.722***
	(36.194)	(36.140)	(0.246)	(0.430)
Farm size (acres)	182.199***	52.076	0.857**	0.669
	(54.510)	(62.527)	(0.346)	(0.549)
Farm size squared (acres)	-29.784	16.562	-0.100	0.283
	(31.305)	(42.938)	(0.213)	(0.364)
Farm size cubed (acres)	1.487	-2.867	-0.001	-0.037
	(4.420)	(5.443)	(0.029)	(0.056)
Farm productive assets (1,000 Ksh)	4.629***	0.904	0.025**	0.008
	(1.560)	(2.110)	(0.012)	(0.017)
Access to credit (dummy)	157.447	164.350	1.117	2.247
	(109.794)	(117.282)	(0.730)	(1.281)
Distance to closest market (km)	10.474	16.544	0.122*	0.156
	(6.480)	(8.905)	(0.056)	(0.099)
Group official (dummy)	100.385	35.229	0.410	0.886
	(68.865)	(110.030)	(0.616)	(0.907)
Surveyed in December (dummy)	140.926	146.200	1.600*	1.497
	(116.033)	(115.739)	(0.841)	(1.169)
Poor agroecology (dummy)	159.997	192.915	1.876	0.170
	(117.296)	(190.591)	(0.987)	(1.590)
Sub-county dummy	Yes	Yes	Yes	Yes
Constant	3764.363***	794.924**	24.831***	21.000***
	(307.369)	(368.374)	(2.328)	(3.544)
Observations	712	712	712	712
R-squared	0.275	0.055	0.185	0.115

Note: Coefficient estimates are shown with robust standard errors clustered at farmer group level in parentheses. All models were estimated with ordinary least squares. AE, male adult equivalent; RE, retinol equivalent. *, **, and *** significant at 10%, 5%, and 1% level, respectively.

Table A3.24. Livestock Commercialization Effects on Purchased Calorie and Nutrient Consumption

Variables	Calories (kcal/day/AE)	Vitamin A (µg RE/day/AE)	Zinc (mg/day/AE)	Iron (mg/day/AE)
Livestock commercialization (0-1)	308.943*	120.867	0.910	3.107**
	(157.776)	(104.713)	(1.016)	(1.351)
Age of household head (years)	-2.581	-2.251	-0.052***	-0.049
	(2.819)	(2.743)	(0.019)	(0.033)
Age squared (years)	-0.042	0.255	-0.000	0.000
	(0.220)	(0.186)	(0.002)	(0.002)
Male household head (dummy)	-133.650	25.592	-1.458*	-0.465
	(127.106)	(43.713)	(0.787)	(0.812)
Education of household head (years)	5.693	2.846	-0.011	0.103
	(11.889)	(5.540)	(0.078)	(0.113)
Household size (AE)	-214.941***	-14.564	-0.835***	-0.888***
	(26.845)	(12.400)	(0.184)	(0.292)
Farm size (acres)	-20.972	-6.507	-0.438	-0.776
	(45.654)	(36.155)	(0.304)	(0.520)
Farm size squared (acres)	-0.256	7.645	0.031	0.336
	(32.209)	(16.217)	(0.196)	(0.287)
Farm size cubed (acres)	0.716	-1.158	0.003	-0.036
	(3.754)	(1.770)	(0.022)	(0.032)
Farm productive assets (1,000 Ksh)	1.486	0.302	0.000	0.016
	(1.605)	(0.832)	(0.009)	(0.010)
Access to credit (dummy)	54.486	43.005	0.542	1.693**
	(85.525)	(32.391)	(0.567)	(0.715)
Distance to closest market (km)	11.013**	3.000	0.097**	0.111
	(4.880)	(3.609)	(0.039)	(0.093)
Group official (dummy)	75.082	78.904*	0.417	0.581
	(79.466)	(40.113)	(0.548)	(0.665)
Surveyed in December (dummy)	137.315	-45.004	1.112	1.390**
	(97.923)	(51.699)	(0.710)	(0.669)
Poor agroecology (dummy)	259.011*	-4.851	2.288**	1.316
	(131.214)	(62.865)	(0.988)	(1.085)
Constant	2593.900***	246.194	14.872***	12.650***
	(226.770)	(173.464)	(1.591)	(2.755)
Sub-county dummies	Yes	Yes	Yes	Yes
Observations	712	712	712	712
R-squared	0.152	0.047	0.111	0.078

Note: Coefficient estimates are shown with robust standard errors clustered at farmer group level in parentheses. All models were estimated with ordinary least squares. AE, male adult equivalent; RE, retinol equivalent. *, **, and *** significant at 10%, 5%, and 1% level, respectively.

Table A3.25. Livestock Commercialization Effects on Own-Produced Calorie and Nutrient Consumption

Variables	Calories (kcal/day/AE)	Vitamin A (µg RE/day/AE)	Zinc (mg/day/AE)	Iron (mg/day/AE)
Livestock commercialization (0-1)	99.016	278.481**	1.386*	2.000
	(127.430)	(132.525)	(0.824)	(1.344)
Age of household head (years)	7.656***	4.117	0.031	0.053
	(2.392)	(4.062)	(0.020)	(0.035)
Age squared (years)	0.230	-0.497**	0.001	-0.003
	(0.197)	(0.189)	(0.002)	(0.002)
Male household head (dummy)	44.705	194.592**	-1.493*	1.347*
	(87.605)	(79.397)	(0.862)	(0.738)
Education of household head (years)	23.243**	2.632	0.156*	0.237*
	(11.061)	(12.108)	(0.085)	(0.126)
Household size (AE)	-143.771***	-69.625**	-0.932***	-0.662**
	(29.226)	(28.569)	(0.237)	(0.318)
Farm size (acres)	224.693***	74.087**	1.341***	1.621***
	(44.622)	(33.394)	(0.336)	(0.450)
Farm size squared (acres)	-40.189	-13.015	-0.184	-0.227
	(26.424)	(30.690)	(0.205)	(0.282)
Farm size cubed (acres)	1.807	0.759	0.002	0.018
	(3.143)	(3.980)	(0.025)	(0.040)
Farm productive assets (1,000 Ksh)	3.511***	0.238	0.026**	-0.004
	(1.262)	(2.295)	(0.010)	(0.016)
Access to credit (dummy)	94.801	141.742	0.684	0.606
	(96.963)	(105.339)	(0.799)	(1.096)
Distance to closest market (km)	-2.836	11.484*	0.020	0.025
	(4.051)	(6.130)	(0.042)	(0.051)
Group official (dummy)	20.554	-58.222	-0.059	0.226
	(83.179)	(90.956)	(0.643)	(0.761)
Surveyed in December (dummy)	-12.573	158.938	0.424	0.045
	(79.007)	(106.277)	(0.592)	(0.893)
Poor agroecology (dummy)	-50.653	228.851	-0.149	-0.713
	(105.090)	(168.798)	(0.947)	(1.308)
Constant	1005.407***	501.250	9.189***	6.834***
	(224.234)	(320.155)	(1.616)	(2.108)
Sub-county dummies	Yes	Yes	Yes	Yes
Observations	712	712	712	712
R-squared	0.215	0.049	0.160	0.089

Note: Coefficient estimates are shown with robust standard errors clustered at farmer group level in parentheses. All models were estimated with ordinary least squares. AE, male adult equivalent; RE, retinol equivalent. *, **, and *** significant at 10%, 5%, and 1% level, respectively.

Table A3.26. Commercialization Effects on Calorie and Nutrient Consumption Using Ward-Level Prices

Variables	Total calories (kcal/day/AE)	Purchased calories (kcal/day/AE)	Own-prod. calories (kcal/day/AE)
Commercialization	699.431***	468.794***	256.567
	(203.101)	(162.128)	(177.064)
	Total vitamin A (µg RE/day/AE)	Purchased vitamin A (µg RE/day/AE)	Own-prod. vit. A (µg RE/day/AE)
Commercialization	129.531	270.247**	-188.204
	(231.493)	(127.763)	(164.811)
	Total zinc (mg/day/AE)	Purchased zinc (mg/day/AE)	Own-prod. zinc (mg/day/AE)
Commercialization	3.537***	2.991***	0.684
	(1.300)	(1.067)	(1.439)
	Total iron (mg/day/AE)	Purchased iron (mg/day/AE)	Own-prod. iron (mg/day/AE)
Commercialization	6.011***	4.050***	1.154
	(2.245)	(1.432)	(1.849)

Note: Coefficient estimates are shown with robust standard errors clustered at farmer group level in parentheses. All models were estimated with ordinary least squares. Only commercialization effects are shown, but for estimation the same explanatory variables as shown in tables 3.4-3.6 in the main paper were included. Ward-level prices instead of total sample mean prices were used to value crop and livestock outputs. AE, male adult equivalent; RE, retinol equivalent. *, **, and *** significant at 10%, 5%, and 1% level, respectively.

Table A3.27. GLM (Fractional Logit) Regression for Estimating Propensity Scores

Variables	GLM Overall commercialization
Age of household head (years)	-0.001
	(0.003)
Male household head (dummy)	0.033
	(0.075)
Education of household head (years)	0.025***
	(0.009)
Household size (AE)	-0.043**
	(0.018)
Farm size (acres)	0.212***
	(0.036)
Farm size squared (acres)	-0.046**
	(0.020)
Farm size cubed (acres)	0.003
	(0.002)
Farm productive assets (1,000 Ksh)	0.002
	(0.001)
Access to credit (dummy)	0.098
	(0.069)
Distance to closest market (km)	-0.004
	(0.005)
Group official (dummy)	0.073
	(0.058)
Surveyed in December (dummy)	0.060
	(0.067)
Poor agroecology (dummy)	-0.244***
	(0.081)
Sub-county dummies	
Sameta	-0.063
	(0.099)
Gucha	0.119
	(0.107)
Kisii Central	-0.222*
	(0.131)
Nyamache	0.072
	(0.099)
Nyamira South	0.046
	(0.110)
Manga	-0.269***
	(0.096)
Masaba North	0.186*
	(0.098)
Constant	-0.604***
	(0.210)
Observations	784
Log pseudo-likelihood	-372.528

Note: Coefficient estimates are shown with robust standard errors in parentheses. Kenyenya was used as reference sub-county. GLM, generalized linear model. *, **, and *** significant at 10%, 5%, and 1% level, respectively.

Table A3.28. Covariate Balancing Tests for Generalized Propensity Score Matching (*t*-Statistics for Mean Differences across Four Treatment Groups)

Covariates	Before matching				After matching			
	TG1[>0,0.30]	TG2 [0.31,0.46]	TG3 [0.47,0.58]	TG4 [0.59,1]	TG1 [>0,0.30]	TG2 [0.31,0.46]	TG3 [0.47,0.58]	TG4 [0.59,1]
Farm and household characteristics								
Age of household head (years)	0.360	-0.076	-1.739*	1.453	0.693	-0.528	-1.586	1.249
Male household head (dummy)	3.354***	0.195	-1.766*	-1.766*	0.176	0.097	-0.727	0.010
Education of household head (years)	5.539***	0.567	-2.762***	-3.274***	1.315	0.227	-1.522	-0.607
Household size (AE)	0.506	-2.323**	1.006	0.803	0.371	-2.067**	0.672	0.519
Farm size (acres)	6.159***	1.616	-2.284**	-5.452***	1.967**	0.264	-0.319	-0.488
Farm size squared (acres)	1.710*	1.470	-0.003	-3.192***	0.832	0.666	0.633	-1.205
Farm size cubed (acres)	1.629	1.383	-0.413	-2.606***	0.646	0.807	-0.166	-0.972
Farm productive assets (1,000 Ksh)	2.913***	0.767	-1.223	-2.451**	0.189	-0.127	-0.017	-0.647
Access to credit (dummy)	2.474**	-1.360	-0.352	-0.755	0.833	-1.467	0.047	0.071
Distance to closest market (km)	-0.177	-1.006	0.491	0.691	0.617	-0.461	0.124	0.422
Group official (dummy)	1.848*	0.815	-0.557	-2.107**	-0.394	0.791	-0.097	-0.492
Surveyed in December (dummy)	1.381	-0.644	1.012	-1.750*	0.332	-0.885	1.232	-0.475
Poor agroecology (dummy)	-1.103	0.122	-1.840*	2.830***	0.302	0.639	-2.836***	1.392
Sub-county dummies								
Sameta	0.741	-0.741	-0.247	0.247	0.223	-0.955	-0.536	0.442
Gucha	0.930	1.179	0.434	-2.553***	-1.327	1.272	1.148	-1.427
Kisii Central	-1.278	-0.255	-0.255	1.790*	0.117	-0.006	-0.888	0.144
Nyamache	1.103	0.857	-1.103	-0.857	-0.357	0.450	-0.803	0.473
Nyamira South	2.231**	-3.051***	0.202	0.607	1.532	-2.708***	0.310	1.007
Manga	-5.766***	0.513	1.956**	3.204***	-1.069	0.571	0.874	0.120
Masaba North	2.634***	0.071	-1.064	-1.633	0.968	-0.563	-0.151	0.406

Note: TG, treatment group. The treatment groups (TG1-TG4) are of equal size, based on the households' level of overall commercialization; levels of commercialization are shown in brackets. Mean values of each TG are compared with mean values of all other TGs combined. *, **, and *** denote significant difference at the 10%, 5%, and 1% level, respectively.

4 Commercialization of the Small Farm Sector and Multidimensional Poverty

Abstract

Global poverty rates have declined considerably, but the number of people living in extreme poverty remains high. Many of the poor are smallholder farmers. Agricultural commercialization – meaning a shift from subsistence to more market-oriented farming – can play a central role in improving smallholder welfare. Previous studies evaluated the impact of agricultural commercialization on income poverty, but whether income gains from commercialization are really used for satisfying basic needs was hardly analyzed up till now. Here, we evaluate the effect of commercialization on income poverty, as well as on the multidimensional poverty index that looks at deprivations in terms of education, nutrition, health, and other dimensions of living standard. Using data from 805 farm households in Kenya, we estimate average treatment effects. We also analyze impact heterogeneity with quantile regressions. Results show that commercialization significantly reduces both income poverty and multidimensional poverty. The magnitude of the income gains is positively correlated with income level, meaning that special market-linkage support for marginalized farms may be required to avoid rising income inequality. However, the effect in terms of reducing basic needs deprivations is strongest among the poorest households, suggesting that agricultural commercialization contributes effectively to achieving the sustainable development goals.

Key words: Agricultural commercialization; Welfare, Multidimensional poverty, Kenya

JEL codes: C21, I32, Q12, Q13

4.1 Introduction

Global poverty rates have declined considerably over the last few decades, but the number of people still living in extreme poverty – below 1.90 US dollars a day – remains high (World Bank 2016). Hence, eradication of poverty continues to be a top priority on the international development agenda (United Nations 2016). Many of the world's poor are smallholder farmers who depend on agriculture as the main source of food, income, and employment. Against this background, agricultural development has been acknowledged as one of the main pathways for poverty alleviation (World Bank 2015; Hazell et al. 2010; de Janvry and Sadoulet 2009).

Commercialization of smallholder agriculture – meaning a shift from subsistence to more market-oriented farming – can lead to productivity growth, income growth, employment growth, and poverty reduction (Bellemare and Novak 2017; Carletto, Corral, and Guefi 2017; von Braun and Kennedy 1994; Barrett 2008). Agricultural commercialization also improves food supply in urban areas, with broader growth and welfare effects. Previous studies confirmed that commercialized farms have higher household incomes than subsistence-oriented farms, after controlling for other relevant factors (von Braun 1995; Tipraqsa and Schreinemachers 2009). A few studies also showed that commercialization contributes to poverty reduction among African smallholders (Muricho et al. 2017; Muriithi and Matz 2015; Olwande et al. 2015).

However, existing studies on poverty effects of commercialization only looked at income poverty. While income (or expenditure) data are widely used to analyze poverty, they cannot fully capture the multidimensional nature of poverty, including deprivations in education, health, nutrition, and other dimensions of living standard. The simple assumption that additional income earned from agricultural commercialization will automatically be spent on satisfying basic needs may not always be true. Different types of income may be controlled by different persons within the farm household and used for different purposes (Meemken, Spielman, and Qaim 2017; von Braun and Kennedy 1994).[14] We contribute to the literature by analyzing the impact of agricultural commercialization on multidimensional poverty, using

[14] A few studies have analyzed the effects of commercialization on nutrition (Carletto et al. 2017; Ogutu, Gödecke, and Qaim 2017), but not on other dimensions of basic needs and living standard.

the multidimensional poverty index described by Alkire and Santos (2014).[15] For comparison, we also analyze the impact of commercialization on income poverty.

The empirical research is based on data from a survey of smallholder farmers in Kenya. As is typical for sub-Saharan Africa, smallholder farmers in Kenya account for the lion's share of total agricultural output and for a large fraction of the population living below the poverty line (World Bank 2017; Wiesmann et al. 2016; Olwande et al. 2015; Mathenge et al. 2014). For the impact analysis, we compare farmers with different levels of commercialization, using a control function approach with instruments to address issues of endogeneity. We estimate average treatment effects of commercialization, as well as heterogeneous treatment effects with quantile regressions. Heterogeneous effects can occur when certain types of households benefit more from commercialization than others. This is important to understand with a view to avoiding rising inequality.

4.2 Data and Measurement of Key Variables

4.2.1 Farm Household Survey

We use data from a farm household survey that we conducted between October and December 2015 in Kisii and Nyamira counties in the western parts of Kenya. These two counties were purposively selected due to the small farm sizes, relatively high poverty rates, diverse agricultural production, and poor road and market infrastructure (Wiesmann et al. 2016; Kisii County Government 2013; Nyamira County Government 2013). Farmers in the study area grow a large number of different crops, such as maize, beans, cassava, sweet potatoes, banana, and vegetables, mostly for home consumption, and to a lesser extent for local market sales. Cash crops such as coffee, tea, and sugarcane are also grown to a limited extent. Many farmers in Kisii and Nyamira are also involved in small-scale livestock keeping, including poultry, small ruminants, and cattle.

As recent census data were not available, we exploited the fact that many of the local farmers are organized in farmer groups or self-help groups for randomly selecting households for the survey. Farmer and self-help groups are registered with the Ministry of Gender, Children, and Social Development. Building on Ministry registries and with support from Africa Harvest, a

[15] In a recent study, Ayuya et al. (2015) used the multidimensional poverty index to analyze the impact of organic farming on smallholder poverty. This is different from the agricultural commercialization question pursued here.

non-governmental organization working in the region, a list of all active groups in Kisii and Nyamira was constructed. From this list, we excluded a few groups that had received specific development support during the last two years in order to avoid any sampling bias. From the remaining groups, we randomly selected 48 groups for our survey (32 groups in Kisii and 16 groups in Nyamira county). In each of these groups, complete member lists were compiled, from which 15-20 households were randomly selected, depending on group size. This resulted in a sample of 824 farm households, spread over 8 different sub-counties and 26 wards.

Data from each household were collected through face-to-face interviews with the household head or sometimes with the spouse. Interviews were carried out in local languages by a team of interviewers, who were trained and supervised by the researchers. The structured questionnaire was carefully tested prior to the survey and included sections on household demographics, agricultural production and marketing, other economic activities of the household, and a large range of institutional and contextual characteristics. Due to missing data, some of the household observations had to be excluded. The sample for this analysis includes 805 households for which complete data are available.

4.2.2 Measuring Agricultural Commercialization

We measure commercialization based on farmers' agricultural production and marketing activities over the 12-month period prior to the survey. We consider all crop and livestock enterprises of the farm household. While semi-subsistence farming is commonplace in the study region, there are hardly any households in the sample who did not sell at least small quantities of their harvest. Hence, measuring commercialization with a simple dummy variable would not be very useful. Instead, we compute the level of commercialization as the share of total farm output sold, a continuous indicator ranging between zero and one. The same approach was also used in previous studies on the effects of commercialization (Carletto et al. 2017; Ogutu et al. 2017; von Braun and Kennedy 1994).

Farmers in Kisii and Nyamira sell their harvest in different types of markets. Small quantities are typically sold to traders at the farm gate or in local village markets. Larger quantities are often sold in the more distant main agricultural markets. Tea and coffee are often delivered to special collection centers at fixed prices. Fixed price arrangements do not exist for food crops in the study region. To calculate the level of commercialization, we use sample average prices for each commodity to value sold and unsold farm output.

4.2.3 Measuring Income Poverty

To analyze the effect of agricultural commercialization on household income and income poverty, we use 12-month data on income from all farm and off-farm economic activities. Farm income is calculated as the value of all agricultural output (sold or unsold) minus production costs. Off-farm income includes the income from all employed and self-employed activities of household members and any transfers and land and capital rents. We report annual household income on a per capita basis expressed in Kenyan shillings (Ksh).

To evaluate effects of commercialization on income poverty, we build on the Foster, Greer, and Thorbecke (1984) class of poverty indicators. We convert per capita income in Ksh to international dollars, using the purchasing power parity (PPP) exchange rate.[16] We define "income poverty" as a dummy variable that takes a value of one if a household's per capita income falls below the international poverty line of 1.90 US dollars a day, and zero otherwise. We also calculate an income poverty gap as follows:

$$y_i = \frac{z - v_i}{z} \tag{10}$$

where z is the poverty line, and v_i is per capita income of a household. Households with incomes above the poverty line are automatically assigned a zero value. The income poverty gap is a continuous variable ranging between zero and one.

4.2.4 Measuring Multidimensional Poverty

Unlike income poverty, which is an indirect approach to assess a household's ability to satisfy basic needs, the multidimensional poverty index (MPI) tries to assess directly whether or not different types of basic needs are actually satisfied. The MPI was proposed by Alkire and Santos (2014). We closely follow their approach and adjust it to the data available in our sample of farm households in Kenya. Adjustments to fit the local context are recommended in the literature (OPHI 2017; Ayuya et al. 2015; Alkire and Santos 2014)

The MPI measures acute poverty by capturing information on the proportion of households within a given population that experience multiple deprivations (multidimensional headcount ratio), and the intensity of their deprivation relative to minimum international standards of well-being. Alkire and Santos (2014) propose three dimensions of poverty – education, health,

[16] In 2015, the PPP exchange rate was 1 US dollar = Ksh 43.89, while the official market exchange rate was 1 US dollar = Ksh 96.30.

and living standard – and 10 indicators for which deprivations are assessed. We use the same dimensions and indicators as proposed by Alkire and Santos (2014), except for three modifications. The first modification is that we do not use the education indicator "no household member has completed 5 years of education", as 99% of our sample have at least one member with 5 or more years of education. We replace this indicator with "the household head has less than 5 years of education". The second and third modifications are that we do not use the health indicators "any child has died in the family" and "any child or adult is malnourished", as we do not have suitable individual-level health and nutrition data. Instead, we use household-level calorie consumption and dietary diversity scores. Descriptions of all 10 indicators used in this study with the corresponding cutoffs are shown in table 4.1.

Using the zero and one values for each of the 10 indicators, we calculate different MPI measures for each sample household. First, we calculate the "total household deprivation score" by summing up the weighted values for each of the 10 indicators, using weights as shown in table 4.1. The total household deprivation score ranges between zero and one, with larger values indicating higher levels of deprivation. Second, we create a "multidimensional poverty dummy", which takes a value of one if a household's total deprivation score is equal to or larger than a certain threshold, and zero otherwise. We use the common threshold of 0.33 (Alkire and Santos 2014). The logic behind this MPI dummy is that a household is considered multidimensionally poor only if it suffers from deprivations in terms of several indicators. Third, we create the "multidimensional poverty intensity", which is equal to the deprivation score if the household is multidimensionally poor (MPI dummy = 1), and zero otherwise.[17] The interpretation of the MPI intensity is similar to the poverty gap, as it measures the magnitude of household deprivations relative to a poverty threshold.

We will use all three MPI measures to evaluate the effects of agricultural commercialization on MPI poverty. A relevant question in this context is to what extent can we actually expect possible income gains from commercialization to affect the different MPI dimensions and indicators? For the indicator related to the household head's level of education an effect can hardly be expected, because adult individuals are unlikely to return to school when their income increases. However, for most of the other indicators related to child education, nutrition, housing conditions, and asset ownership (table 4.1) changes through income gains and other possible effects of commercialization are plausible.

[17] Thus, the MPI intensity can either be 0 or take values in the range between 0.33 and 1. For an individual household it cannot take value between 0 and 0.33, even though the sample mean value can be in this range when taking the average across all households, including MPI poor and non-poor.

Table 4.1. Dimensions and Indicators of the Multidimensional Poverty Index

Dimension and indicator	Description and deprivation cutoff	Relative weight
Education		
Years of schooling	The household head has less than 5 years of education	1/6
Child school attendance	The household has a school-aged child not attending school up to class 8	1/6
Health		
Nutrition 1	The household consumes less than 2400 kcal per day and adult male equivalent (AE)	1/6
Nutrition 2	The household has a dietary diversity score of 5 or less out of 10 possible food groups [a]	1/6
Living standard		
Electricity	The household has no electricity	1/18
Sanitation	The household's toilet facility is not improved, or it is improved but shared with other households	1/18
Drinking water	The household does not have access to safe drinking water	1/18
Floor	The household has dirt, sand, or dung floor	1/18
Cooking fuel	The household cooks with dung, wood or charcoal	1/18
Asset ownership	The household does not own more than one radio, TV, telephone, bike, motorbike or refrigerator and does not own a car or truck	1/18

Notes: The indicators are very similar to those in Alkire and Santos (2014), except for small modifications in three indicators (years of schooling, nutrition 1, nutrition 2) as explained in the text. [a] The 10 food groups used are those recommended for the minimum dietary diversity score for women (FAO and FHI 360 2016).

4.3 Estimation Strategy

4.3.1 Conditional Mean Effects

To determine the effects of commercialization on income poverty and MPI poverty, we estimate the following regression:

$$y_i = \alpha_0 + \alpha_1 C_i + \alpha_2 \boldsymbol{X}_i + \varepsilon_i \tag{11}$$

where y_i is the poverty indicator for household i, C_i is the level of commercialization, $\boldsymbol{X}_i$ is a vector of control variables, and ε_i is a random error term, clustered at farmer group level. We estimate separate models for each of the different poverty indicators (see previous subsection), always controlling for relevant household, farm, and contextual variables that may influence poverty through pathways other than commercialization. For the models with continuous dependent variables (income, poverty gap, deprivation scores, MPI intensity), we use ordinary least squares (OLS) estimators. Some of these variables are censored at zero and one, so that we also use fractional logit estimators as robustness checks (Papke and Wooldridge 1996). For the models with binary dependent variables (income poverty dummy, MPI dummy), we use logit estimators.

The main coefficient of interest in equation (11) is α_1, which measures the effect of commercialization on household income or poverty. We expect a positive coefficient α_1 when using absolute household income as the dependent variable, and a negative coefficient α_1 when using the poverty indicators. In other words, we expect commercialization to have income-increasing and poverty-reducing effects. However, the level of commercialization is potentially endogenous, which would lead to correlation between C_i and ε_i and biased estimates of α_1. Endogeneity of C_i may arise from unobserved heterogeneity, reverse causality, or measurement error. We test and control for endogeneity bias with a control function approach and instrumental variables, as explained below.

4.3.2 Control Function Approach

We use a control function (CF) approach (Wooldridge 2015; Smith and Blundell 1986; Rivers and Voung 1988) to account for potential endogeneity of the commercialization variable C_i. The CF approach uses instrumental variables (IV) for proper identification of causal effects and is more flexible with respect to functional form than standard IV estimators, such as two-stage least squares. Our choice of the CF approach is motivated by the fact that the commercialization variable is bounded between zero and one (and has nontrivial zero observations), leading to non-linear corner solutions. In such cases, the CF approach is more efficient than two-stage least squares (Verkaart et al. 2017; Wooldridge 2015).

The CF approach involves predicting residuals from a first-stage regression model of the determinants of commercialization, which must include one or more valid instruments. We use a fractional logit estimator for this first-stage regression. The predicted residuals are then included as an additional covariate in the second-stage regression – the income or poverty model in equation (11). A significant coefficient of the residual term in equation (11) would mean that C_i is endogenous. In that case, including the residual term would correct for endogeneity bias of the coefficient α_1. However, an insignificant residual term would mean that the null hypothesis of exogeneity of C_i cannot be rejected. In that case, excluding the residual term would produce unbiased and more efficient estimates.

As indicated, the CF approach requires one or more valid instruments in the first-stage regression. For an instrument to be valid it has to be correlated with the level of commercialization C_i but must not affect income or poverty outcomes y_i through other mechanisms. We use two instruments, namely the average number of motorcycles owned by households living in the same ward as the farmer himself/herself, and the average number of main market sellers in the ward. In Kenya, a ward is an administrative unit that is larger than a village, but smaller than a sub-county. As explained above, the farm survey covered farm households in 26 different wards. On average, 31 households were interviewed in each ward. The two instruments are explained and tested for validity in the following.

The first instrument – the average number of motorcycles in the ward – is constructed by counting the number of motorcycles owned by sample households in each ward (excluding the farmer himself/herself), and then dividing by the number of sample households in the ward. Less than 10% of the households own any motorized means of transportation (average number of motorcycles in the sample is 0.08). Yet, the markets are often distant, so that it is difficult for farmers to make larger sales of agricultural output without using a motor vehicle. Since most of the feeder roads in the study area are not paved and public transport services barely exist, motorcycles owners tend to provide transport services to households located in the same area. Farmers often use these transport services, as do local traders who buy farm produce at the farm gate and sell in more distant markets. Thus, more motorcycles in the ward imply better market access. The average number of motorcycles in the ward is significantly correlated with the level of commercialization (*p*-value=0.007). Hence, the first condition for instrument validity is satisfied.

To test for the second condition of instrument validity we need to show that the number of motorcycles does not affect income and poverty through mechanisms other than commercialization. Since we use the average number of motorcycles owned by households in the ward, as opposed to individual household ownership, the instrument is not significantly associated with any of the household-level poverty indicators, with or without controlling for other possible poverty determinants (table A4.1 in the appendix). We also tested for possible correlations between the instrument and other farm and household-level characteristics, as transport services may possibly change households' access to information, inputs, and technologies. None of the correlation coefficients was statistically significant (table A4.2 in the appendix). The correlation between the number of motorcycles in the ward and other ward-level wealth indicators were also insignificant (table A4.3). These test results suggest that the second condition of instrument validity is also satisfied.

The second instrument – the average number of main market sellers in the ward – is constructed by counting the number of farmers in each ward who sold at least some of their produce in main agricultural markets (excluding the farmer himself/herself) and then dividing by the number of sample households in the ward. As mentioned above, the main agricultural markets are the locations where larger quantities of agricultural output are typically sold (smaller quantities are sold at the farm gate or in local village markets). Hence, farmers who sell some or all of their produce in the more distant main markets are likely to have a higher level of commercialization. One-third of the farm households in our sample sell at least some of their harvest in main agricultural markets. As expected, these farmers have significantly larger farm output and sales revenues than their colleagues not selling in the main agricultural markets (table A4.4).

However, why should the presence of main market sellers in the ward affect the level of commercialization of other farmers? The choice of this instrument is inspired by the recent strand of literature on peer learning, showing that farmers tend to learn about the benefits of innovations from their peers (Magnan et al. 2015; Krishnan and Patnam 2013). We posit that farmers in the same ward will likely belong to the same social networks. Hence, farmers who benefit from selling in main agricultural markets may potentially influence their peers to also supply such markets, leading to higher levels of commercialization. Farmers living in the same neighborhood may also benefit from collective action, which can help to reduce transaction costs and enhance market participation (Fischer and Qaim 2012). Andersson et al. (2015) used data from Kenya to show that farmers whose neighbors supplied supermarkets were more likely to also supply supermarkets, due to joint organization and shared transport costs. In our data, we find that the average number of main market sellers in the ward is significantly correlated with the degree of commercialization of individual farmers (*p*-value=0.000). Hence, the first condition of instrument validity is satisfied.

But is the number of main market sellers in the ward also affecting income or poverty outcomes directly? This could happen when more commercialized and better-off households cluster in the same wards. However, such clustering does not seem to occur in the study region. The instrument is not correlated with any of the ward-level wealth indicators, as shown in table A4.3 in the appendix. Nor do we find significant correlation between the instrument and individual farm household characteristics (table A4.2). When correlating the number of main market sellers in the ward with household-level poverty indicators, some of the correlation coefficients are statistically significant. However, once we control for commercialization in regression models the instrument coefficients turn insignificant (table A4.1 in the appendix). Hence, there does not seem to be any effect of the instrument on income or poverty through mechanisms other than commercialization, thus the second condition for instrument validity is also satisfied.

We also tested for overidentifying restrictions with both instruments, as shown in table A4.5 in the appendix. Based on the test results we cannot reject the null hypothesis of joint instrument exogeneity. Hence, we conclude that the two instruments are valid.

4.3.3 Heterogeneous Treatment Effects

The effects of commercialization on household income and poverty may be heterogeneous, meaning that some households may benefit more than others. From a social development perspective, we are particularly interested in understanding whether the poorest households benefit to the same extent as the relatively richer ones. The model in equation (11) estimates average treatment effects, but it cannot estimate impact heterogeneity. We use quantile regressions to examine potential impact heterogeneity of agricultural commercialization. Quantile regressions allow one to examine whether the effect of a particular regressor changes over the conditional distribution of the dependent variable, instead of only analyzing the regressor’s average effect (Koenker and Hallock 2001; Buchinsky 1998).

The conditional quantile functions of the income and poverty indicators (y_i) given regressor x_i (in our case the level of commercialization, C_i) can be expressed as follows:

$$y_i = x_i'\beta_\tau + \upsilon_{\tau_i}, \qquad Q_\tau(y_i \mid x_i) = x_i'\beta_\tau, \tag{12}$$

where $Q_\tau(y_i \mid x_i)$ is the conditional quantile of y_i at quantile τ , with $0 < \tau < 1$. β_τ is the vector of parameters to be estimated. The parameters are obtained by minimizing the following equation, which is solved by linear programming:

$$\min_{\beta_\tau} \frac{1}{n} \left\{ \sum_{i: y_i \geq x_i'\beta_\tau} \tau \,|\, y_i - x_i'\beta_\tau \,| + \sum_{i: y_i < x_i'\beta_\tau} (1-\tau) \,|\, y_i - x_i'\beta_\tau \,| \right\} \tag{13}$$

Equation (13) implies that the parameters can be estimated at different points or quantiles (τ) of the dependent variable by minimizing the sum of asymmetrically weighted absolute residuals (Koenker and Hallock 2001).

We estimate quantile regressions for key continuous outcome variables – namely per capita income, multidimensional poverty intensity, and total household deprivation scores – to evaluate potential effects of commercialization on inequality. Effects of commercialization are estimated at five different quantiles ($\tau = 0.10$, 0.25, 0.50, 0.75, and 0.90). We use the same variables in equation (11) as regressors. For interpretation of the effects of C_i it is important to consider the distribution of the dependent variable. When using absolute income as dependent variable, $\tau = 0.10$ represents the poorest group of households. When using the MPI intensity and total deprivation scores as dependent variables, $\tau = 0.10$ represents the least-poor households.

4.4 Results and Discussion

4.4.1 Descriptive Statistics

Table 4.2 shows summary statistics for the full sample of farm households and also disaggregated by level of commercialization. For these descriptive comparisons we subdivide the sample into quartiles according to the household level of commercialization and compare the most commercialized (highest quartile – MC25%) with the least commercialized (lowest quartile – LC25%) households.

Table 4.2. Summary Statistics by Level of Commercialization

Variables	Full sample Mean	MC25% Mean	LC25% Mean	Mean difference
Socioeconomic characteristics				
Commercialization (share of farm output sold, 0-1)	0.44	0.70	0.16	0.55***
	(0.21)	(0.09)	(0.09)	
Age of household head (years)	49.27	48.35	48.34	0.01
	(12.57)	(11.22)	(13.63)	
Male household head (dummy)	0.77	0.82	0.67	0.15***
	(0.42)	(0.39)	(0.47)	
Education of household head (years)	8.94	9.69	7.80	1.89***
	(3.77)	(3.19)	(4.09)	
Household size (adult equivalents)	3.99	3.92	3.89	0.03
	(1.58)	(1.62)	(1.63)	
Farm size (acres)	1.61	2.04	1.14	0.90***
	(1.27)	(1.55)	(0.95)	
Farm productive assets (1,000 Ksh)	19.93	23.78	15.54	8.24***
	(23.69)	(25.43)	(20.84)	
Household income (1,000 Ksh/year)	180.53	281.36	90.69	190.67***
	(218.46)	(285.81)	(103.12)	
Off-farm income (dummy)	0.81	0.78	0.81	-0.04
	(0.39)	(0.42)	(0.39)	
Access to credit (dummy)	0.78	0.80	0.69	0.11**
	(0.41)	(0.40)	(0.46)	
Distance to closest market (km)	4.91	4.60	4.97	-0.37
	(7.01)	(5.25)	(7.53)	
Distance to closest extension agent (km)	4.34	3.89	5.52	-1.63***
	(4.93)	(4.67)	(5.40)	
Household head/spouse is a group official (dummy)	0.35	0.41	0.28	0.13***
	(0.48)	(0.49)	(0.45)	
Poor agroecology [a] (dummy)	0.13	0.07	0.16	0.09***
	(0.34)	(0.26)	(0.37)	
Farm production diversity (no. of food crop/livestock species)	11.11	11.21	10.33	0.88**
	(4.39)	(4.72)	(4.06)	
Livestock ownership (tropical livestock units - TLU)	1.73	1.60	1.41	0.19
	(1.62)	(1.65)	(1.42)	
Motorcycles in ward [b] (number)	0.08	0.10	0.08	0.02***
	(0.05)	(0.06)	(0.05)	
Main market sellers in ward [b] (number)	0.32	0.36	0.29	0.07***
	(0.11)	(0.12)	(0.11)	
Farm productivity and input use				
Value of crop output (1,000 Ksh/acre)	75.81	105.13	70.32	34.80***
	(81.94)	(110.42)	(97.12)	
Seed expenditure (Ksh/acre)	3184.90	3212.07	3018.04	194.03
	(3892.72)	(3792.63)	(2411.09)	
Fertilizer expenditure (Ksh/acre)	6269.29	6569.09	5383.40	1185.69**
	(5479.26)	(6338.84)	(4515.33)	
Manure expenditure (Ksh/acre)	708.89	666.33	608.87	57.46
	(2958.03)	(2794.36)	(2171.11)	
Pesticide expenditure (Ksh/acre)	659.72	911.25	330.46	580.79***
	(1626.87)	(2038.22)	(1080.75)	
Observations	805	201	202	

Notes: Standard deviations are shown in parentheses. MC25%, 25% most commercialized households; LC25%, 25% least commercialized households; Ksh, Kenyan shillings. [a] Variable takes a value of one if a farmer reported serious crop loss due to pests and diseases. [b] Ward-level variables were divided by the number of households interviewed in each ward to allow meaningful comparison. *, **, and *** significant at 10%, 5%, and 1% level, respectively.

On average, sample households sell 44% of their farm output, while the most and least commercialized quartiles sell 70% and 16% of their farm output, respectively (table 4.2). As one would expect, more commercialized households tend to have larger farm sizes, higher levels of education, more assets, and better access to credit and extension. Commercialization is also positively associated with several other socioeconomic variables, as well as with farm input use and productivity.

Table 4.3 presents the share of households deprived in each of the ten MPI indicators. A large variation across the different indicators is observed. While a relatively small share of the sample households is deprived in the education indicators, most of the households are deprived in access to electricity (89%) and modern cooking fuel (97%). Figure 4.1 confirms that deprivations are much more prevalent for the living standard indicators than for the education and health indicators. Figure 4.1 also shows that the least-commercialized households are significantly more deprived than the most-commercialized households in terms of all the three MPI dimensions.

Table 4.3. Share of Households Deprived in MPI Indicators (Indicators Ranked by Share of Deprived Households)

Indicator	Deprivations cut-off	Full sample	MC25%	LC25%	Mean difference
Years of schooling	Household head has less than 5 years of education (dummy)	0.130	0.070	0.203	-0.133***
		(0.337)	(0.255)	(0.403)	
Nutrition 2 (dietary quality)	Household consumed 5 or less out of 10 possible food groups (dummy)	0.142	0.090	0.248	-0.158***
		(0.349)	(0.286)	(0.433)	
Child school attendance	Household has a school-aged child not attending up to class 8 (dummy)	0.154	0.144	0.178	-0.033
		(0.361	(0.352)	(0.384)	
Nutrition 1 (calorie intake)	Household consumes less than 2400 kcal/day/AE (dummy)	0.266	0.179	0.337	-0.157***
		(0.442)	(0.384)	(0.473)	
Asset ownership	Household does not own more than one of specified assets [a] (dummy)	0.338	0.279	0.436	-0.157***
		(0.473)	(0.449	(0.497)	
Sanitation	Household's toilet facility is not improved (dummy)	0.553	0.488	0.649	-0.161***
		(0.498)	(0.487)	(0.479)	
Drinking water	Household does not have access to safe drinking water (dummy)	0.557	0.532	0.633	-0.101**
		(0.497)	(0.500)	(0.483)	
Floor	Household has dirt, sand or dung floor (dummy)	0.737	0.711	0.787	-0.076*
		(0.441	(0.454)	(0.410)	
Electricity	Household has no electricity (dummy)	0.889	0.861	0.911	-0.050
		(0.314)	(0.347)	(0.286)	
Cooking fuel	Household cooks with dung, wood or charcoal (dummy)	0.968)	0.955	0.985	-0.030*
		(0.177)	(0.207)	(0.121)	
	Observations	805	201	202	

Note: Standard deviations are shown in parentheses. For further details of indicator descriptions, see table 4.1. MC25%, 25% most commercialized households, LC25%, 25% least commercialized households, AE, male adult equivalent. *, **, and *** significant at 10%, 5%, and 1% level, respectively.

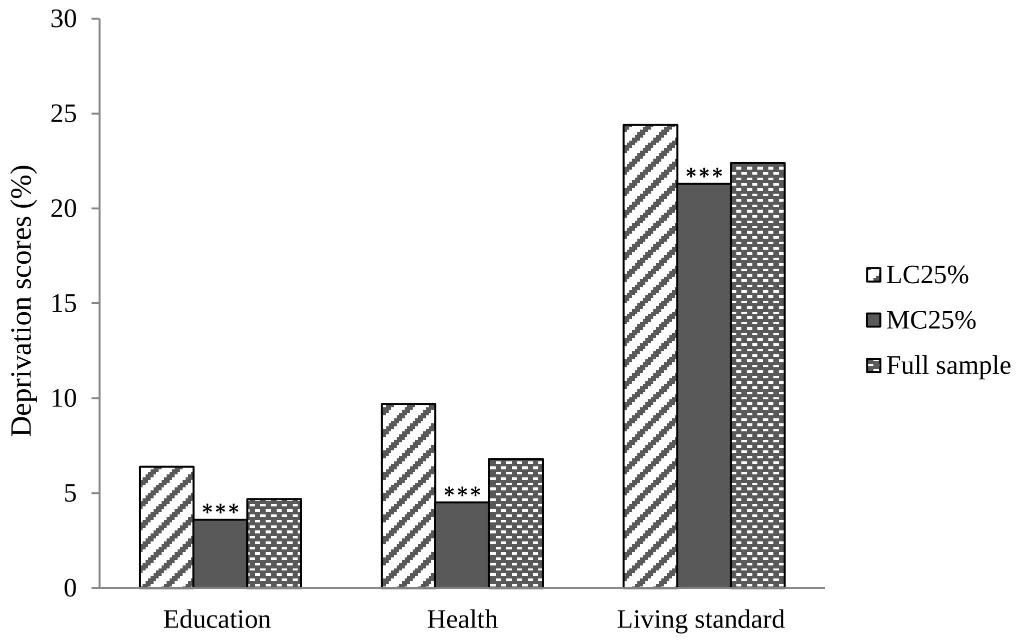

Figure 4.1. Mean Household Deprivation Scores by Multidimensional Poverty Dimension.

Notes: Deprivation scores in each of the three dimensions can range between 0 and 33%. MC25%, 25% most commercialized households; LC25%, 25% least commercialized households *** Difference between most and least commercialized households is significant at 1% level.

Table 4.4 presents summary statistics for the different poverty indicators. Sixty-two percent of the households are poor in terms of income poverty, meaning that they have less than 1.90 US dollars per capita and day in PPP terms. The income poverty headcount is much larger among the least-commercialized than among the most-commercialized households. The income poverty gap is also much larger among the least-commercialized households.

Table 4.4. Poverty Indicators by Level of Commercialization

Variable	Full sample Mean	MC25% Mean	LC25% Mean	Mean Difference
Household income (1,000 Ksh)	180.53	281.36	90.69	190.67***
	(218.46)	(285.81)	(103.12)	
Per capita income (1,000 Ksh)	35.09	54.38	17.29	37.09***
	(44.90)	(59.92)	(19.26)	
Income poverty (dummy)	0.62	0.40	0.83	-0.43***
	(0.49)	(0.49)	(0.38)	
Income poverty gap (0-1)	0.34	0.16	0.56	-0.40***
	(0.34)	(0.25)	(0.34)	
Total household deprivation score (0-1)	0.34	0.29	0.40	-0.11***
	(0.17)	(0.14)	(0.18)	
Multidimensional poverty (dummy)	0.51	0.39	0.65	-0.26***
	(0.50)	(0.49)	(0.48)	
Multidimensional poverty intensity (0-1)	0.24	0.17	0.33	-0.16***
	(0.25)	(0.22)	(0.27)	
Observations	805	201	202	

Notes: Standard deviations are shown in parentheses. MC25%, 25% most commercialized households; LC25%, 25% least commercialized households; Ksh, Kenyan shillings; *, **, and *** significant at 10%, 5%, and 1% level, respectively.

In terms of multidimensional poverty, the mean total deprivation score of 0.34 implies that the average household suffers from 34% of the possible deprivations. As explained earlier in section 4.2, a household is classified as MPI poor when the total household deprivation score is equal or greater than 0.33; this applies to 51% of the households. The mean MPI intensity is 0.24 across all households.[18] Table 4.4 shows that the least-commercialized farm households are significantly more affected by the prevalence and depth of MPI poverty than the most commercialized households.

While the comparisons between more and less commercialized households are in line with our hypothesis that commercialization contributes to poverty reduction, the differences in table 4.4 cannot be interpreted as causal effects, because they do not control for possible confounding factors. We control for confounding factors in the following subsections through the regression models explained in section 4.3.

4.4.2 Conditional Mean Effects

The first-stage results of the CF approach are shown in table A4.6 in the appendix. The residuals from this first-stage regression are included in the CF models in tables 4.5-4.7. The residual term is insignificant in all the models, so we cannot reject the null hypothesis of commercialization being exogenous. Hence, we prefer the models without the residual terms for interpretation, because they produce more efficient estimates. However, we show both versions of the models. The signs and magnitudes of the estimated commercialization coefficients are similar with and without the residual terms included, which underlines the robustness of the general findings.

Table 4.5 presents the effects of commercialization on per capita income. Commercialization has a positive and significant effect. The level of commercialization is a continuous variable ranging from zero to one, which has to be taken into account when interpreting the coefficient magnitudes. The estimate in column (1) of table 4.5 suggests that a 0.1 increase (10 percentage points) in the level of commercialization increases annual per capita income by 5,000 Ksh, which is equivalent to a 14% gain relative to the total sample mean income. Relative to the lower mean income of the least commercialized households, 5,000 Ksh of additional income would represent a gain of 29%. A hypothetical shift from a zero level of commercialization to 44% – the sample mean level of commercialization – would more than

[18] The MPI intensity calculated here is very similar to the MPI intensity of 0.25 reported in a recent study for rural Kenya in general (OPHI 2017).

double per capita income. These are net income gains of commercialization after controlling for other important factors that can also influence income such as education, farm size, ownership of other productive assets, agroecology, as well as infrastructure and institutional conditions.

Table 4.5. Effect of Commercialization on Per capita Income

	Per capita income (1,000 Ksh)		Log of per capita income	
	(1)	(2)	(3)	(4)
Variables	OLS	CF	OLS	CF
Commercialization (0-1)	50.124***	52.855***	1.688***	1.712***
	(8.448)	(9.508)	(0.143)	(0.196)
Age of household head (years)	-0.087	-0.097	-0.002	-0.002
	(0.143)	(0.134)	(0.002)	(0.002)
Age squared (years)	0.008	0.010	-0.000	-0.000
	(0.010)	(0.008)	(0.000)	(0.000)
Male household head (dummy)	8.454***	8.259***	0.324***	0.323***
	(2.932)	(2.764)	(0.079)	(0.067)
Education of household head (years)	1.530***	1.468***	0.042***	0.042***
	(0.366)	(0.370)	(0.009)	(0.009)
Household size (number)	-5.570***	-5.482***	-0.132***	-0.131***
	(1.050)	(1.080)	(0.014)	(0.017)
Farm size (acres)	3.328	2.550	0.124***	0.117**
	(1.824)	(3.444)	(0.037)	(0.054)
Farm size squared (acres)	0.154	0.351	0.000	0.002
	(0.781)	(1.562)	(0.018)	(0.022)
Farm size cubed (acres)	0.105	0.089	0.000	-0.000
	(0.127)	(0.296)	(0.002)	(0.003)
Farm productive assets (1,000 Ksh)	0.486***	0.479***	0.009***	0.009***
	(0.094)	(0.084)	(0.001)	(0.001)
Access to credit (dummy)	2.786	2.329	0.150***	0.146**
	(3.276)	(3.141)	(0.049)	(0.068)
Distance to closest market (km)	0.013	0.026	0.001	0.001
	(0.271)	(0.252)	(0.005)	(0.004)
Group official (dummy)	-1.278	-1.538	-0.011	-0.013
	(2.945)	(2.957)	(0.054)	(0.053)
Off-farm income (dummy)	18.511***	18.492***	0.797***	0.797***
	(3.316)	(3.110)	(0.072)	(0.066)
Poor agroecology (dummy)	2.724	3.517	0.111	0.118
	(3.349)	(4.419)	(0.070)	(0.085)
Livestock ownership (TLU)	1.111	1.224	0.073***	0.074***
	(1.014)	(0.959)	(0.018)	(0.017)
Residual from first stage		-3.174		-0.029
		(8.985)		(0.167)
Sub-county dummies	Yes	Yes	Yes	Yes
Constant	-9.212	-6.521	1.225***	1.250***
	(9.483)	(10.465)	(0.199)	(0.266)
Observations	805	805	805	805
R-squared	0.366	0.366	0.567	0.567

Notes: Coefficient estimates are shown with robust standard errors in parentheses. In columns (1) and (2), standard errors are clustered at farmer group level. In columns (2) and (4), standard errors are bootstrapped with 1000 replications. OLS, ordinary least squares; CF, control function estimator; TLU, tropical livestock units. *, **, and *** significant at 10%, 5%, and 1% level, respectively.

Given the large standard deviation of per capita income, we also estimated the same model with a log-transformation of the dependent variable. Results are shown in columns (3) and (4) of table 4.5. This alternative specification confirms the large positive income effects of agricultural commercialization. The estimates suggest that a 10 percentage point increase in the level of commercialization would increase per capita income by 17% after controlling for other factors.

Table 4.6. Effect of Commercialization on Income Poverty

	Income poverty (dummy)		Income poverty gap (0-1)	
	(1)	(2)	(3)	(4)
Variables	Logit	CF	OLS	CF
Commercialization (0-1)	-0.506***	-0.573***	-0.531***	-0.562***
	(0.072)	(0.110)	(0.049)	(0.075)
Age of household head (years)	0.001	0.002	0.001	0.001
	(0.001)	(0.001)	(0.001)	(0.001)
Age squared (years)	0.000	0.000	0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)
Male household head (dummy)	-0.179***	-0.175***	-0.127***	-0.125***
	(0.052)	(0.041)	(0.031)	(0.025)
Education of household head (years)	-0.014***	-0.013**	-0.013***	-0.012***
	(0.004)	(0.005)	(0.003)	(0.003)
Household size (number)	0.071***	0.069***	0.041***	0.040***
	(0.007)	(0.008)	(0.005)	(0.005)
Farm size (acres)	-0.045***	-0.025	-0.036**	-0.027
	(0.019)	(0.030)	(0.014)	(0.020)
Farm size squared (acres)	0.000	-0.005	0.001	-0.001
	(0.012)	(0.016)	(0.007)	(0.008)
Farm size cubed (acres)	-0.001	-0.001	0.000	0.000
	(0.002)	(0.003)	(0.001)	(0.001)
Farm productive assets (1,000 Ksh)	-0.004***	-0.004***	-0.002***	-0.002***
	(0.001)	(0.001)	(0.000)	(0.000)
Access to credit (dummy)	-0.063**	-0.051	-0.067***	-0.062**
	(0.030)	(0.040)	(0.019)	(0.025)
Distance to closest market (km)	0.001	0.001	0.000	-0.000
	(0.002)	(0.002)	(0.001)	(0.001)
Group official (dummy)	-0.007	-0.001	0.000	0.003
	(0.032)	(0.031)	(0.020)	(0.020)
Off-farm income (dummy)	-0.292***	-0.291***	-0.227***	-0.227***
	(0.036)	(0.044)	(0.020)	(0.023)
Poor agroecology (dummy)	-0.003	-0.022	-0.029	-0.038
	(0.050)	(0.054)	(0.026)	(0.031)
Livestock ownership (TLU)	-0.018*	-0.020*	-0.023***	-0.025***
	(0.010)	(0.010)	(0.005)	(0.006)
Residual from first stage		0.079		0.036
		(0.096)		(0.064)
Sub-county dummies	Yes	Yes	Yes	Yes
Constant			0.905***	0.874***
			(0.067)	(0.093)
Observations	805	805	805	805
(Pseudo) R-squared	0.342	0.343	0.472	0.472

Notes: In columns (1) and (2), average partial effects are shown with robust standard errors in parentheses. In columns (3) and (4), coefficient estimates are shown with robust standard errors in parentheses. In columns (1) and (3), standard errors are clustered at farmer group level. In columns (2) and (4), standard errors are bootstrapped with 1000 replications. OLS, ordinary least squares; CF, control function estimator; TLU, tropical livestock units. *, **, and *** significant at 10%, 5%, and 1% level, respectively.

Table 4.6 shows the effects of commercialization on income poverty. The average partial effect estimate of -0.506 in column (1) implies that full commercialization would halve the probability of falling below the poverty line of 1.90 US dollars a day. A 10 percentage point increase in the level of commercialization would reduce the prevalence of income poverty by 5.1 percentage points. Columns (3) and (4) of table 4.6 show that commercialization also reduces the income poverty gap significantly. Holding other factors constant, a 10 percentage point increase in the level of commercialization reduces the poverty gap by an average of 5.3 percentage points. This is equivalent to a 16% reduction in the mean poverty gap of the total sample. Fractional logit specifications of the poverty gap model are shown in table A4.7 in the appendix with very similar results.

Table 4.7 shows the effects of commercialization on multidimensional poverty. The average partial effect estimate of -0.226 in column (1) implies that full commercialization would reduce the probability of being MPI poor by 22.6 percentage points. A 10 percentage point increase in the level of commercialization reduces the prevalence of MPI poverty by 2.3 percentage points. Columns (3) and (4) of table 4.7 show that commercialization also reduces the multidimensional poverty intensity. A 10 percentage point increase in the level of commercialization reduces the MPI intensity by approximately 1.5 percentage points. Fractional logit specifications of the MPI intensity model are shown in table A4.7 in the appendix with very similar results.

Table 4.7. Effect of Commercialization on Multidimensional Poverty

	Multidimensional poverty (dummy)		Multidimensional poverty intensity (0-1)	
	(1)	(2)	(3)	(4)
Variables	Logit	CF	OLS	CF
Commercialization (0-1)	-0.226***	-0.189	-0.153***	-0.144***
	(0.083)	(0.124)	(0.042)	(0.057)
Age of household head (years)	-0.001	-0.001	-0.000	-0.000
	(0.001)	(0.002)	(0.001)	(0.001)
Age squared (years)	0.000	-0.000	-0.000	-0.000
	(0.000)	(0.000)	(0.000)	(0.000)
Male household head (dummy)	-0.028	-0.031	-0.022	-0.023
	(0.038)	(0.043)	(0.022)	(0.022)
Education of household head (years)	-0.040***	-0.041***	-0.024***	-0.024***
	(0.004)	(0.005)	(0.002)	(0.003)
Household size (number)	0.019**	0.021**	0.014***	0.014***
	(0.009)	(0.009)	(0.005)	(0.005)
Farm size (acres)	-0.027	-0.038	-0.015	-0.018
	(0.019)	(0.035)	(0.009)	(0.016)
Farm size squared (acres)	-0.006	-0.004	-0.001	-0.000
	(0.010)	(0.017)	(0.005)	(0.007)
Farm size cubed (acres)	0.002	0.002	0.001	0.001
	(0.001)	(0.003)	(0.001)	(0.001)
Farm productive assets (1,000 Ksh)	-0.004***	-0.004***	-0.002***	-0.002***
	(0.001)	(0.001)	(0.000)	(0.000)
Access to credit (dummy)	-0.102***	-0.108***	-0.049***	-0.050**
	(0.029)	(0.041)	(0.016)	(0.020)
Distance to closest market (km)	-0.001	-0.001	-0.001*	-0.001
	(0.002)	(0.002)	(0.001)	(0.001)
Group official (dummy)	-0.014	-0.017	-0.024	-0.024
	(0.031)	(0.037)	(0.016)	(0.017)
Off-farm income (dummy)	-0.049	-0.050	-0.031	-0.031
	(0.053)	(0.044)	(0.025)	(0.020)
Poor agroecology (dummy)	0.013	0.023	0.013	0.015
	(0.051)	(0.057)	(0.026)	(0.027)
Livestock ownership (TLU)	-0.018	-0.017	-0.013***	-0.012**
	(0.011)	(0.013)	(0.005)	(0.006)
Residual from first stage		-0.044		-0.010
		(0.108)		(0.051)
Sub-county dummies	Yes	Yes	Yes	Yes
Constant			0.670***	0.678***
			(0.053)	(0.073)
Observations	805	805	805	805
(Pseudo) R-squared	0.199	0.199	0.300	0.301

Notes: In columns (1) and (2), average partial effects are shown with robust standard errors in parentheses. In columns (3) and (4), coefficient estimates are shown with robust standard errors in parentheses. In columns (1) and (3), standard errors are clustered at farmer group level. In columns (2) and (4), standard errors are bootstrapped with 1000 replications. OLS, ordinary least squares; CF, control function estimator; TLU, tropical livestock units. *, **, and *** significant at 10%, 5%, and 1% level, respectively.

4.4.3 Mechanisms of Poverty Reduction

The results in tables 4.6 and 4.7 suggest that the effects of commercialization on multidimensional poverty are smaller than the effects on income poverty. This is not surprising. Income poverty falls automatically when poor households experience income gains that are sufficiently large to lift them above the income poverty line. However, whether the additional income is really used to satisfy basic needs is a question that cannot be answered with income-related poverty indicators alone.

The results with the multidimensional poverty indicators in table 4.7 suggest that the additional income from commercialization is indeed used to satisfy basic needs to a significant extent. In other words, agricultural commercialization contributes to poverty reduction regardless of whether poverty is assessed and measured through indirect or direct approaches. As one would expect, the multidimensional poverty effects differ by MPI dimension, as is shown in table A4.8 in the appendix. This also explains why the MPI effects are smaller than the income poverty effects. Commercialization has a small decreasing effect on education deprivations, but this effect is not statistically significant. As discussed above, education deprivations among sample households are relatively small anyway, and the education level of the household head will hardly change through additional commercialization income.

The commercialization effect on living standard deprivations is somewhat larger and statistically significant (table A4.8). While some of the living standard indicators – such as housing conditions, cooking fuel and asset ownership – can easily be improved when the income increases, other indicators – such as access to electricity and safe drinking water – may require broader infrastructure investments that are beyond the scope of individual households. The largest effects of commercialization on MPI poverty are observed in terms of reducing health deprivations. As explained, the indicators used for the health dimension are calorie consumption and dietary quality, which households can improve through rising incomes. Given widespread food insecurity among smallholder farm households, the finding that commercialization improves nutrition is certainly welcome.

So far, we have assumed that the effects of commercialization on multidimensional poverty are primarily channeled through the income pathway. This is confirmed in table A4.9 in the appendix, where we regress the MPI intensity on income and other explanatory variables. Income gains contribute significantly to reducing MPI intensity. Interestingly, the effect is stronger for farm income than for total household income. Tables A4.10 and A4.11 in the appendix show some of the main pathways how commercialization contributes to rising farm and total household incomes, namely through production increases resulting from higher input intensity and productivity. Table A4.11 also shows that agricultural productivity and the value of production significantly contribute to income poverty and multidimensional poverty reduction.

4.4 Heterogeneous Treatment Effects

We now examine whether the effects of commercialization on income and multidimensional poverty are heterogeneous. As mentioned, it is possible that different types of households benefit more or less than others. We estimate heterogeneous treatment effects with quantile regressions. The estimation results are shown in tables A4.12-A4.14 in the appendix. The commercialization effects are shown graphically in figure 4.2.

Panel (A) of figure 4.2 shows the quantile effects of commercialization on per capita income. With per capita income as dependent variable, the 0.10 quantile includes the poorest, whereas the 0.90 quantile includes the richest sample households. As can be seen, commercialization has significant positive effects on per capita income across all quantiles. However, the absolute income gains for the poorest households are smaller than those for the richest households. This difference between the lowest and highest quantile is statistically significant (table 4.8). Hence, commercialization increases income inequality.

Table 4.8. Wald Test for Equality of Quantile Coefficients (Conditional Slope Parameters)

Poverty indicator	Wald test F-statistic of $\tau = 0.90$ versus…	
	$\tau = 0.10$	$\tau = 0.50$
Per capita income (Ksh 1,000)	5.60^{**}	1.68
Multidimensional poverty intensity (0-1)	-	0.44
Total household deprivation score (0-1)	2.78^{*}	0.55

Notes: * and **significant at 10% and 5% level, respectively.

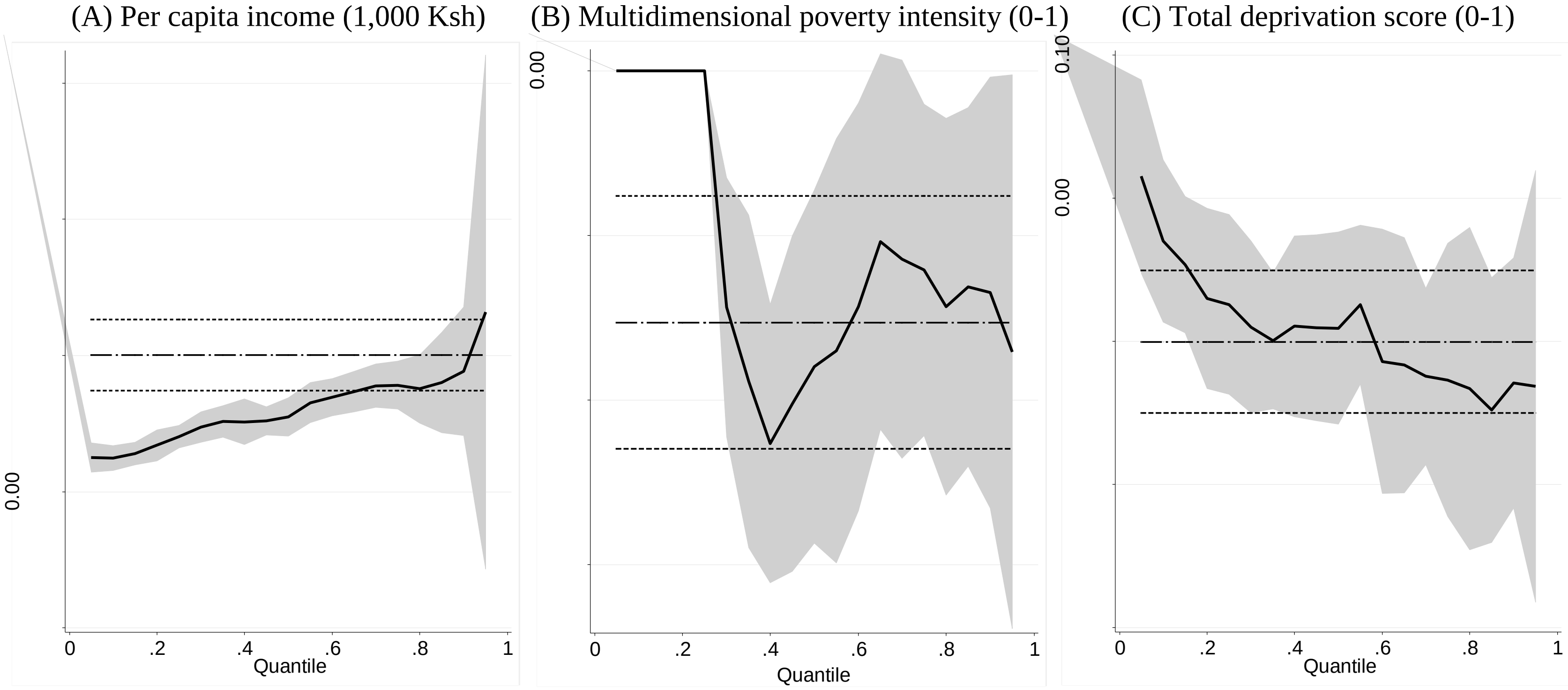

Figure 4.2. Quantile Regression Estimates for Per capita Income, Multidimensional Poverty Intensity, and Total Household Deprivation Score.

Notes: Conditional quantile estimates are shown with thick solid lines. Shaded areas indicate 95% confidence intervals. Dashed-dotted horizontal lines show point estimates from ordinary least square models. Dotted horizontal lines show 95% confidence intervals from ordinary least square models. Details of the estimation results are shown in tables A4.12-A4.14 in the appendix.

Panel (B) of figure 4.2 depicts the quantile effects for multidimensional poverty intensity. Here it is important to stress that larger values of the dependent variable indicate higher levels of poverty, so the quantile interpretation is reversed: the lowest quantile represents the better-off households, meaning those least affected by multidimensional poverty. As can be seen, commercialization significantly reduces MPI intensity for the poorer households in the higher quantiles. Although some variation occurs, the differences between the effects for these upper quantiles are not statistically significant (table 4.8). For the lower quantiles, effects could not be estimated, because the better-off households have an MPI intensity of zero.

However, many of the households not classified as MPI poor still suffer from deprivations in terms of individual indicators. Therefore, we also estimated a quantile regression using the total household deprivation score as dependent variable. Results are shown in panel (C) of figure 4.2. Again, the lowest quantile represents the better-off households, meaning those least affected by the different deprivations. As can be seen, commercialization significantly reduces total household deprivations across all quantiles, except for the richest households (0.10 quantile). The effects are stronger for the poorest households, and the difference between the highest and lowest quantile is statistically significant (table 4.8). These results suggest that – in spite of rising income inequality – agricultural commercialization effectively contributes to satisfying basic needs, especially among the most deprived farm households.

4.5. Conclusion

Using data from smallholder farm households in Kenya and various regression techniques, we have analyzed the effects of agricultural commercialization on household income, income poverty, and multidimensional poverty. The contribution to the literature lies particularly in the analysis of impacts on multidimensional poverty. Looking at various dimensions of poverty, as we have done using the multidimensional poverty index (MPI) proposed by Alkire and Santos (2014), is important, because it cannot simply be assumed that income gains from commercialization will always be spent on satisfying basic needs. The MPI captures three dimensions of poverty, namely education, health/nutrition, and living standard, each with various indicators. Another novelty of our study is that we have estimated heterogeneous treatment effects of commercialization using quantile regressions, which has not been done previously.

Results showed that commercialization increases per capita income in smallholder farm households and reduces income poverty and multidimensional poverty. Even though the effects are significant for all the outcome variables, the impact on income poverty is stronger than the impact on multidimensional poverty. This is plausible because some of the basic needs deprivations can be remedied more easily than others. For instance, households can improve their nutrition and housing conditions when their income increases, but may depend on public infrastructure investments before they can notably improve their access to electricity and safe drinking water. Hence, impact evaluations based on income poverty measures alone may overestimate reductions in terms of various household deprivations.

The quantile regression results showed that absolute gains in per capita income through commercialization are larger for the richer than for the poorer households, suggesting that commercialization contributes to rising income inequality. However, we did not find heterogeneous effects of commercialization on the multidimensional poverty intensity. For reductions in total household deprivations we even found stronger effects for the most deprived households. We conclude that agricultural commercialization is an important and effective mechanism towards achieving the sustainable development goals.

An important policy implication is that commercialization can be fostered through enhancing smallholder market access in terms of investments in road and market infrastructure and strengthening relevant market institutions. Market-linkage support specifically tailored to the needs of the poor may potentially also help to avoid rising income inequality. However, commercialization alone will not suffice to eradicate multidimensional poverty in the small farm sector. Complementary interventions to improve access to sanitation, healthcare, drinking water, education, and sustainable energy will be required such that rising household demand for these basic goods and services resulting from income gains is effectively met by high-quality supply.

While our results proved to be robust across different model specifications, two limitations should briefly be discussed. First, we relied on cross-section observational data which means that dealing with possible endogeneity is challenging. Follow-up research with panel data could further improve the identification strategy and could also provide interesting insights into possible longer-term effects of commercialization. Second, the concrete results from smallholder farmers in Kenya should not be generalized. The situation of farmers in the study area is typical for the African small farm sector, so that some broader general lessons can be learned. But in terms of the specific effects of commercialization on different MPI indicators, results may differ by geographical context.

Appendix A4

Table A4.1. Association between Instruments and Poverty Indicators

Poverty indicators	Motorcycles in ward		Main market sellers in ward	
	Correlation coefficient	Regression coefficient	Correlation coefficient	Regression coefficient
Household income (1,000 Ksh)	0.040 (0.257)	48.045 (0.585)	0.073 (0.039)	21.024 (0.609)
Per capita income (1,000 Ksh)	-0.011 (0.749)	-40.478 (0.215)	0.033 (0.349)	-6.540 (0.662)
Income poverty (dummy)	-0.037 (0.289)	-1.106 (0.600)	-0.092 (0.010)	-0.865 (0.300)
Household poverty gap (0-1)	-0.043 (0.224)	-0.290 (0.775)	-0.112 (0.001)	-0.274 (0.573)
Multidimensional poverty (dummy)	-0.006 (0.866)	1.163 (0.544)	-0.035 (0.316)	0.533 (0.581)
Multidimensional poverty intensity (0-1)	-0.029 (0.405)	0.023 (0.886)	-0.058 (0.098)	0.017 (0.847)

Notes: The average number of motorcycles and of main market sellers in the ward are used as instruments for commercialization. *p*-values are shown in parentheses. The regression coefficients were estimated with models that include the instruments plus all other explanatory variables as those used in tables 4.5-4.7 of the main paper.

Table A4.2. Correlation between Instruments and Farm Household Characteristics

Variables	Motorcycles in ward		Main market sellers in ward	
	Correlation coefficients	*p*-value	Correlation coefficients	*p*-value
Household nutrition knowledge score [a]	0.032	0.355	-0.022	0.536
Household seed expenditure per acre	-0.015	0.668	-0.047	0.184
Household fertilizer expenditure per acre	-0.025	0.477	-0.019	0.589
Household pesticide expenditure per acre	-0.057	0.106	-0.054	0.129
Household manure expenditure per acre	0.018	0.605	0.001	0.973

Notes: The average number of motorcycles and main market sellers in ward are used as instruments for commercialization. [a] Household nutrition knowledge was computed based on four questions related to knowledge of (1) food sources of rich in vitamin A; (2) vitamin A deficiencies; (3) food sources rich in iron and (4) iron deficiencies.

Table A4.3. Correlation between Instruments and Mean Socioeconomic Characteristics at Ward Levels

Variables	Motorcycles in ward		Main market sellers in ward	
	Correlation coefficients	*p*-value	Correlation coefficients	*p*-value
Mean education of household head (years)	0.054	0.794	0.137	0.505
Mean household income (1,000 Ksh)	0.038	0.852	0.164	0.424
Mean farm productive assets (1,000 Ksh)	-0.039	0.851	0.083	0.687
Mean farm size (acres)	0.036	0.860	0.221	0.278

Notes: The average number of motorcycles and main market sellers in ward are used as instruments for commercialization. Socioeconomic characteristics were computed by averaging across all sample households in the ward.

Table A4.4. Mean Differences in Output and Sales between Main Market Sellers and Non-sellers

Variables	Full sample	Participants	Nonparticipants	Mean difference
Value of output (1,000 Ksh)	139.382	178.152	120.990	57.162***
	(176.251)	(240.692)	(131.716)	
Value of sales (1,000 Ksh)	71.976	102.937	57.289	45.648***
	(108.139)	(146.002)	(80.552)	
Value of inputs (1,000 Ksh)	13.798	16.842	12.354	4.488***
	(14.120)	(17.214)	(12.138)	

Notes: Standard deviations are shown in parentheses. Ksh, Kenyan shillings; 1 US dollar = 96.3 Ksh. *, **, and *** significant at 10%, 5%, and 1% level, respectively.

Table A4.5. Over-identification Tests for Joint Instrument Exogeneity with Different Poverty Indicators

Variables	*p*-value
Multidimensional poverty (dummy)	0.288
Multidimensional poverty index (1-0)	0.777
Income poverty (dummy)	0.103
Household poverty gap (1-0)	0.777
Per capita income (Ksh 1,000)	0.120
Log of per capita income	0.526

Note: Based on the insignificant *p*-values we fail to reject over-identification test and conclude that the instruments are jointly valid or exogenous.

Table A4.6. First-Stage Regression Model for Determinants of Commercialization

Variables	GLM (fractional logit) Commercialization
Main market traders in ward (number)	2.314***
	(0.448)
Motorcycles in ward (number)	-2.448***
	(0.901)
Age of household head (years)	0.003
	(0.003)
Age squared (years)	-0.000**
	(0.000)
Male household head (dummy)	0.019
	(0.059)
Education of household head (years)	0.020**
	(0.009)
Household size (number)	-0.030*
	(0.016)
Farm size (acres)	0.239***
	(0.037)
Farm size squared (acres)	-0.064***
	(0.019)
Farm size cubed (acres)	0.005**
	(0.002)
Farm productive assets (1,000 Ksh)	0.002*
	(0.001)
Access to credit (dummy)	0.166**
	(0.082)
Distance to closest market (km)	-0.003
	(0.004)
Group official (dummy)	0.087*
	(0.053)
Off-farm income (dummy)	0.019
	(0.069)
Poor agroecology (dummy)	-0.219***
	(0.075)
Livestock ownership (TLU)	-0.034*
	(0.019)
Constant	-1.325***
	(0.247)
Sub-county dummies	Yes
Log pseudo-likelihood	-379.534
p-values showing instrument relevance	
p-value of motorcycles in ward=0.007	
p-value of main market sellers in ward=0.000	
p-value of excluded instruments (joint significance)=0.000	
Observations	805

Note: GLM, generalized linear model. Coefficient estimates are shown with robust standard errors clustered at farmer group level in parentheses *, **, and *** significant at 10%, 5%, and 1% level, respectively. TLU, tropical livestock units. We observe a negative coefficient on motorcycles in ward due to its high correlation with main market sellers in ward. When separate regressions are run with each of the instruments the coefficients are positive as expected.

Table A4.7. Effect of Commercialization on Income Poverty Gap and Multidimensional Poverty Intensity, Estimated with Fractional Logit Models

	Income poverty gap (0-1)		Multidimensional poverty intensity (0-1)	
	(1)	(2)	(3)	(4)
Variables	Fractional logit	CF	Fractional logit	CF
Commercialization (0-1)	-0.481***	-0.507***	-0.129***	-0.127**
	(0.047)	(0.071)	(0.039)	(0.057)
Age of household head (years)	0.001	0.001	-0.001	-0.001
	(0.001)	(0.001)	(0.001)	(0.001)
Age squared (years)	0.000	0.000	-0.000*	-0.000*
	(0.000)	(0.000)	(0.000)	(0.000)
Male household head (dummy)	-0.108***	-0.106***	-0.007	-0.007
	(0.027)	(0.023)	(0.019)	(0.020)
Education of household head (years)	-0.013***	-0.012***	-0.023***	-0.023***
	(0.003)	(0.003)	(0.002)	(0.002)
Household size (number)	0.043***	0.042***	0.014***	0.014***
	(0.005)	(0.005)	(0.004)	(0.004)
Farm size (acres)	-0.033***	-0.025	-0.014	-0.015
	(0.013)	(0.020)	(0.004)	(0.016)
Farm size squared (acres)	0.001	0.000	-0.003	-0.003
	(0.007)	(0.009)	(0.005)	(0.007)
Farm size cubed (acres)	-0.001	-0.001	0.001	0.001
	(0.001)	(0.002)	(0.001)	(0.001)
Farm productive assets (1,000 Ksh)	-0.003***	-0.003***	-0.003***	-0.003***
	(0.001)	(0.001)	(0.001)	(0.000)
Access to credit (dummy)	-0.061***	-0.056***	-0.045***	-0.029**
	(0.017)	(0.024)	(0.014)	(0.014)
Distance to closest market (km)	0.000	0.000	-0.002*	-0.001**
	(0.001)	(0.001)	(0.001)	(0.001)
Group official (dummy)	0.004	0.006	-0.021	-0.036***
	(0.020)	(0.020)	(0.016)	(0.011)
Off-farm income (dummy)	-0.213***	-0.213***	-0.031	-0.024*
	(0.017)	(0.021)	(0.023)	(0.013)
Poor agroecology (dummy)	-0.027	-0.035	0.012	0.006
	(0.026)	(0.030)	(0.024)	(0.019)
Livestock ownership (TLU)	-0.025***	-0.026***	-0.013**	-0.011***
	(0.007)	(0.007)	(0.006)	(0.004)
Residual from first stage		0.030		-0.002
		(0.063)		(0.050)
Sub-county dummies	Yes	Yes	Yes	Yes
Observations	805	805	805	805
Log pseudo-likelihood	-320.825	-320.826	-305.255	-305.254

Notes: Average partial effects are shown with robust standard errors in parentheses. In columns (1) and (3), standard errors are clustered at farmer group level. In columns (2) and (4), standard errors are bootstrapped with 1000 replications. CF, control function estimator; TLU, tropical livestock units. *, **, and *** significant at 10%, 5%, and 1% level, respectively.

Table A4.8. Effect of Commercialization on Different Multidimensional Poverty Dimensions

Variables	Total deprivations in education	Total deprivations in health	Total deprivations in living standards
Commercialization	-0.007	-0.058***	-0.025**
	(0.011)	(0.018)	(0.013)
Control variables	Yes	Yes	Yes
Log-pseudolikelihood	-107.411	-149.372	-292.941
Observations	805	805	805

Notes: Average partial effects are shown with robust standard errors in parentheses. The dependent variables are deprivation scores in each of the three dimensions, all three ranging between 0 and 0.33. The same explanatory variables as used in tables 4.5-4.7 of the main paper were used for estimation but are not shown here for brevity. ** and *** significant at 5% and 1% level, respectively.

Table A4.9. Effect of Income on Multidimensional Poverty Intensity

Variables	(1) MPI	(2) MPI	(3) MPI
Per capita income (1,000 Ksh)	-0.008***		
	(0.000)		
Household income (1,000 Ksh)		-0.001***	
		(0.000)	
Farm income (1,000 Ksh)			-0.003***
			(0.001)
Control variables	Yes	Yes	Yes
Constant	1.271***	1.102***	1.174***
	(0.276)	(0.273)	(0.271)
Log pseudo-likelihood	-304.689	-305.098	-304.546
Observations	805	805	805

Notes: Coefficient estimates are shown with robust standard errors clustered at farmer group level in parentheses. All models were estimated with a fractional logit estimator. The same explanatory variables as used in tables 4.5-4.7 of the main paper were used for estimation but are not shown here for brevity. *, **, and *** significant at 10%, 5%, and 1% level, respectively.

Table A4.10. Effect of Commercialization on Farm input Use and Land Productivity

Variable	Seed cost per acre	Fertilizer cost per acre	Manure cost per acre	Pesticide cost per acre	Value of output per acre
Commercialization	1874.632***	4400.213***	1917.692***	1199.029***	68.752***
	(624.347)	(1035.559)	(454.973)	(422.123)	(22.885)
Constant	2093.498*	4096.084***	1218.896**	414.260	55.766***
	(1133.162)	(1400.210)	(585.913)	(333.402)	(19.763)
Sub-county dummies	Yes	Yes	Yes	Yes	Yes
Observations	805	805	805	805	805
R-squared	0.104	0.144	0.098	0.064	0.101

Notes: Coefficient estimates are shown with robust standard errors clustered at farmer group level in parentheses. All models estimated with OLS except for the manure model, which was estimated with a control function estimator (bootstrapped standard errors with 1000 replications), due to commercialization being endogenous in the manure model. The same explanatory variables as used in tables 4.5-4.7 of the main paper were used for estimation but are not shown here for brevity. *, **, and *** significant at 10%, 5%, and 1% level, respectively.

Table A4.11. Associations between Farm Input, Output, and Poverty

Variables	(1) Per capita income	(2) Log per capita income	(3) Income poverty gap (0-1)	(4) MPI intensity (0-1)
Total value of inputs (1,000 Ksh)	0.608***	0.015***	-0.006***	-0.002***
	(0.136)	(0.002)	(0.001)	(0.000)
Total value of output (1,000 Ksh)	0.103***	0.002***	-0.002***	-0.001**
	(0.018)	(0.003)	(0.000)	(0.000)
Value of output per acre (1,000 Ksh)	0.095***	0.003***	-0.002***	-0.001*
	(0.022)	(0.001)	(0.000)	(0.000)

Notes: Coefficient estimates are shown with robust standard errors clustered at farmer group level in parentheses. Each coefficient was estimated with a separate model. Models in columns (1) and (2) estimated with ordinary least squares. Models in columns (3) and (4) estimated with fractional logit. In all models, the same explanatory variables as used in tables 4.5-4.7 of the main paper were used for estimation but are not shown here for brevity. *, **, and *** significant at 10%, 5%, and 1% level, respectively.

Table A4.12. Quantile Regression for Per capita Income (1,000 Ksh)

Variables	OLS	Quantile 0.10	Quantile 0.25	Quantile 0.50	Quantile 0.75	Quantile 0.90
Commercialization (0-1)	50.124***	12.353***†	20.175***†	27.339***†	39.071***	44.172***
	(8.448)	(2.716)	(2.702)	(4.567)	(5.488)	(13.026)
Constant	-9.212	-6.412*	-9.756***	-5.907***	-7.175	12.406
	(9.483)	(3.586)	(4.791)	(5.968)	(8.196)	(18.693)
Control variables	Yes	Yes	Yes	Yes	Yes	Yes
(Pseudo) R-squared	0.366	0.154	0.185	0.248	0.288	0.338

Notes: *N* = 805. OLS coefficient estimates are shown with robust standard errors clustered at farmer group level in parentheses. Quantile regression coefficients are shown with bootstrapped standard errors (1000 replications) in parentheses. The same explanatory variables as used in tables 4.5-4.7 of the main paper were used for estimation but are not shown here for brevity *, **, and *** significant at 10%, 5%, and 1% level, respectively. † coefficient is significantly different from OLS estimate.

Table A4.13. Quantile Regression for Multidimensional Poverty Intensity (0-1)

Variables	OLS	Quantile 0.50	Quantile 0.75	Quantile 0.90
Commercialization (0-1)	-0.153***	-0.180***	-0.121**	-0.135**
	(0.042)	(0.056)	(0.051)	(0.055)
Constant	0.670***	0.727***	0.761***	0.864***
	(0.053)	(0.080)	(0.068)	(0.094)
Control variables	Yes	Yes	Yes	Yes
(Pseudo) R-squared	0.300	0.259	0.185	0.180

Notes: *N* = 805. OLS coefficient estimates are shown with robust standard errors clustered at farmer group level in parentheses. Quantile regression coefficients are shown with bootstrapped standard errors (1000 replications) in parentheses. Regression for the 0.10 and 0.25 quantiles could not be estimated due to a large proportion of zeros for the MPI intensity in these relatively better-off groups. The same explanatory variables as used in tables 4.5-4.7 of the main paper were used for estimation but are not shown here for brevity *, **, and *** significant at 10%, 5%, and 1% level, respectively.

Table A4.14. Quantile Regression for Total Household Deprivation Scores (0-1)

		Quantile				
Variables	OLS	0.10	0.25	0.50	0.75	0.90
Commercialization (0-1)	-0.100***	-0.030†	-0.074***	-0.091***	-0.127***	-0.129**
	(0.025)	(0.035)	(0.033)	(0.034)	(0.039)	(0.051)
Constant	0.628***	0.383***	0.502***	0.610***	0.702***	0.862***
	(0.034	(0.051)	(0.048)	(0.050)	(0.057)	(0.082)
Control variables	Yes	Yes	Yes	Yes	Yes	Yes
(Pseudo) R-squared	0.343	0.169	0.169	0.205	0.217	0.225

Notes: N = 805. OLS coefficient estimates are shown with robust standard errors clustered at farmer group level in parentheses. Quantile regression coefficients are shown with bootstrapped standard errors (1000 replications) in parentheses. The same explanatory variables as used in tables 4.5-4.7 of the main paper were used for estimation but are not shown here for brevity *, **, and *** significant at 10%, 5%, and 1% level, respectively. † coefficient is significantly different from OLS estimate.

5 Conclusion and Policy Implications

5.1 Main Findings

We have examined how agricultural training can be combined with training in nutrition and marketing to increase farmers' adoption of a new bean variety biofortified with iron and zinc. To the best of our knowledge, this is the first study to analyze how improved designs of agricultural extension can contribute to making smallholder farming more nutrition-sensitive. We have also evaluated the effect of commercialization on household food security and dietary quality, with a particular focus on calorie and micronutrient consumption. Besides, we have analyzed the potential channels through which commercialization influences farm household nutrition by examining the role of income, gender, and possible substitution between consumption of own-produced and purchased foods, aspects that were hardly examined by previous studies. Lastly, we have examined the effect of commercialization on multidimensional and income poverty, adding further insights to the literature on the welfare implications of commercialization.

To examine how agricultural training can be combined with training in nutrition and marketing to increase farmers' adoption of a new bean variety biofortified with iron and zinc, we conducted a randomized controlled trial with farmers in Western Kenya. In the experiment, we implemented three extension designs (treatments) and evaluated their effects on the adoption of KK15 beans that were biofortified with iron and zinc. Difference-in-difference estimates showed that intensive training offered by agricultural extension officers and tailored to local conditions can increase technology adoption considerably within a relatively short period of time. In all three treatments, the adoption of KK15 increased from less than 1% before the RCT started to more than 20% one year later. This rapid increase in adoption in the treatment groups suggests that farmers are willing to adopt pro-nutrition technologies, if they are well informed about their attributes and implications, even when the technologies are not primarily designed to increase productivity and income. Although farmers in the RCT received a 30% subsidy on the price of KK15 seeds, they had to pay for the seeds and therefore made a real adoption decision considering expected benefits and costs.

Comparison of the different treatments revealed interesting additional insights. Farmers who had received agricultural training and nutrition training were more likely to adopt KK15 than farmers who had only received agricultural training. Comparison of the treatment-on-the-treated (TOT) effects suggest that additional nutrition training further increased adoption rates by 10-12 percentage points over and above the effects of stand-alone agricultural training. This additional effect of nutrition training may not surprise, because of the positive nutritional attributes of KK15. However, it should be noted that these attributes of KK15 were communicated to farmers in the agricultural training sessions. The nutrition training sessions covered broader aspects related to healthy nutrition, balanced diets, and the health consequences of nutrient deficiencies. It seems that knowledge about these broader nutrition aspects helped farmers to better appreciate the nutrition attributes of KK15, thus resulting in higher adoption rates. The nutrition training may certainly have positive effects on household diets and health beyond KK15 adoption. Analysis of such wider effects is beyond the scope of this dissertation.

As mentioned, we have also evaluated the effect of commercialization on household food security and dietary quality, with a particular focus on calorie and micronutrient consumption. We also examined transmission channels through which commercialization influences nutrition by looking at the role of income, gender, and possible substitution between the consumption of own-produced and purchased foods. The analysis was based on cross-sectional data from 805 farm households in Western Kenya. A control function approach was used to address possible endogeneity problems, and generalized propensity scores to estimate continuous treatment effects. The estimation results showed that commercialization has positive effects on food security and dietary quality. Higher levels of commercialization significantly contribute to improved calorie, zinc, and iron consumption. For vitamin A consumption, the effects of commercialization were insignificant. The positive effects for most dietary indicators are primarily due to rising cash incomes, allowing households to purchase more food from the market. However, rather than substituting for own-produced foods, purchased foods are added to the diet with increasing levels of commercialization. Hence, commercialization contributes to higher levels of dietary diversity. That commercialized households continue to also rely on own-produced foods is probably attributable to persistent market failures. Maintaining a certain level of subsistence is a typical response to market risks and high transaction costs. Only for highly commercialized households, the role of own-produced foods in the diets starts to decline.

Our analysis of how commercialization may affect gender roles within farm households showed that commercialization leads to a higher share of farm revenues being controlled by male household members. This shift from female to male control has negative partial effects on the consumption of calories and micronutrients, especially vitamin A. These results confirm earlier research showing that women tend to spend more on dietary quality than men (Hoddinott and Haddad 1995; Fischer and Qaim 2012a).

Finally, using data from smallholder farm households in Kenya and various regression techniques, we have analyzed the effects of agricultural commercialization on household income, income poverty, and multidimensional poverty. The contribution to the literature lies particularly in the analysis of impacts on multidimensional poverty. Results showed that commercialization increases per capita income in smallholder farm households and reduces income poverty and multidimensional poverty. Even though the effects are significant for all of the outcome variables, the impact on income poverty is stronger than the impact on multidimensional poverty. This is plausible because some of the basic needs deprivations can be remedied more easily than others. For instance, households can improve their nutrition and housing conditions when their income increases, but may depend on public infrastructure investments before they can notably improve their access to electricity and safe drinking water. Hence, impact evaluations based on income poverty measures alone may overestimate reductions in terms of various household deprivations.

The quantile regression results showed that absolute gains in per capita income through commercialization are larger for the richer than for the poorer households, suggesting that commercialization contributes to rising income inequality. However, we did not find heterogeneous effects of commercialization on the multidimensional poverty intensity. For reductions in total household deprivations we even found stronger effects for the most deprived households. We conclude that agricultural commercialization is an important and effective mechanism towards achieving the sustainable development goals.

5.2 Policy Implications

We find that combining agricultural training with nutrition training increases adoption rates over and above the effects of agricultural training alone. This has important policy implications. Nutrition education is usually not delivered through the agricultural extension service, but through specialized nutrition and health workers. Our results suggest that combining agricultural and nutrition training in agricultural extension approaches is possible. Of course, the nutrition training should be designed together with nutrition experts, and the agricultural extension officers first need to be trained before they can effectively deliver nutrition training to farm households. However, the high personnel and logistics cost of reaching out to households in rural areas is a major impediment for more widespread coverage of nutrition and health education campaigns. Based on our results we argue that closer cooperation between agricultural extension and nutrition and health organizations can be a cost-effective way to promote pro-nutrition innovations in smallholder farm households.

We also find that commercialization contributes to improved nutrition – calorie, zinc, and iron consumption – in the small farm sector. An important policy implication is that enhancing market access is a key strategy to make smallholder agriculture more nutrition-sensitive. However, the role of women should receive particular attention. The evidence suggests that women may lose decision-making power with increasing levels of commercialization, but this may possibly be prevented through more gender-sensitive approaches and awareness-building initiatives (Meemken and Qaim 2018). We also stress that commercialization alone will not suffice to address all types of malnutrition. Commercialization helps to increase cash income, but the consumption of certain micronutrients – such as vitamin A – does not seem to be particularly responsive to income growth. Hence, more specific, complementary interventions may be needed.

Finally, we also find that commercialization reduces multidimensional and income poverty in the small farm sector, but it may lead to rising income inequality between more and less commercialized households. An important policy implication is that commercialization can be fostered through enhancing smallholder market access in terms of investments in road and market infrastructure and strengthening relevant market institutions. Market-linkage support specifically tailored to the needs of the poor may potentially also help to avoid rising income inequality. However, commercialization alone will not suffice to eradicate multidimensional poverty in the small farm sector. Complementary interventions to improve access to sanitation, healthcare, drinking water, education, and sustainable energy will be required such that rising household demand for these basic goods and services resulting from income gains is effectively met by high-quality supply.

5.3 Limitations of the Study

The study region in Western Kenya with very small farm sizes, diverse production systems, limited market access due to infrastructure constraints, and relatively high rates of malnutrition is typical of the African small farm sector. Hence, some of the general findings will also be relevant beyond this specific setting. However, in chapter 2, the exact estimates of the treatment effects should not be generalized. There are particularly two factors in our RCT that may possibly reduce the external validity of the empirical estimates. First, our extension treatments were fairly intense. Within a period of nine months, farmers in all treatment groups were offered seven agricultural training sessions. In some of the treatment groups, three nutrition training and three marketing training sessions were offered in addition. Outside an experiment, the training frequency and intensity may be lower, meaning that the effects on technology adoption may be lower too. Second, we only analyzed the short-term adoption effects, as the follow-up survey was carried out less than one year after the treatments had started. Technology adoption is a process over time. Most farmers seemed to be satisfied with KK15 during the first year of adoption, so it is likely that adoption rates will further increase in the future, among both treated and untreated farmers. Further research is needed on how the design of agricultural extension approaches can be improved in order to increase the adoption of pro-nutrition technologies. Our study is only an initial step in this direction.

While several tests confirmed the robustness of our findings in chapter 3, a few limitations remain. First, the analysis relies on cross-sectional data, which limits the strength of the identification strategy. Nor do cross-section data allow the analysis of longer-term effects, which is a drawback because welfare impacts may vary over time (Carletto et al. 2010; Carletto et al. 2011). Follow-up studies with panel data and observed changes in the level of commercialization over time would be very useful. Second, the 7-day food consumption recall data provide a reasonable snapshot of dietary quality at the household level, but they do not account for seasonality and intra-household food distribution. Although we showed that seasonal variations in diets are relatively small and that the household-level nutrition indicators are significantly correlated with individual-level measures, the collection and use of higher-frequency, individual-level nutrition data would be very useful for more detailed analyses. Third, the use of 12-months recall data for farm production and marketing activities is likely associated with certain levels of imprecision. In this respect, higher-frequency data collected in various seasons of a year could reduce possible measurement errors. Fourth, while we tried to analyze possible effects of commercialization on gender roles within the household, a more rigorous analysis of the gender transmission channel would benefit from a larger number of gender-disaggregated variables.

A final limitation of the essay in Chapter 3 is that of external validity of the results. Of course, the concrete results are context-specific and should not simply be generalized. Nevertheless, we argue that some broader lessons can probably be learned. As mentioned, the study region in Western Kenya with very small farm sizes, diverse production systems, limited market access due to infrastructure constraints, and relatively high rates of malnutrition is typical of the African small farm sector. Hence, some of the general findings will also be relevant beyond this specific setting. One characteristic of the study region in Western Kenya that is more location-specific is the fact that seasonality in agricultural production and consumption is not very pronounced. This is related to ample rainfall in various months of each year. Effects of commercialization may be different in regions with stronger seasonality and higher risk of drought. The fact that our sample was drawn from households that are organized in farmer groups should be mentioned, but is unlikely to reduce external validity in a significant way. We focused on farmer groups because this allowed us to randomly sample from existing lists in the absence of county and village census data. According to our own field observations, the households that are organized in farmer groups are not notably different from other farm households living in the study region.

In chapter 4, while our results proved to be robust across different model specifications, two limitations are noteworthy. First, some possible endogeneity issues remain since we rely on cross-sectional data. Follow-up research with panel data would provide more insights on the longer-term effects of commercialization. Second, the concrete results from smallholder farmers in Kenya should not be generalized. The situation of farmers in the study area is typical for the African small farm sector, so that some broader general lessons can be learned. But in terms of the specific effects of commercialization on different MPI indicators, results may differ by geographical context.

References

Alkire, S., and M.E., Santos. 2014. Measuring Acute Poverty in the Developing World: Robustness and Scope of the Multidimensional Poverty Index. *World Development 59*: 251-274.

Anderson, J.R., and G. Feder. 2007. Agricultural Extension. *Handbook of Agricultural Economics 3*: 2343-2378.

Anderson, J.R., and G. Feder. 2004. Agricultural Extension: Good Intentions and Hard Realities. *The World Bank Research Observer 19*(1): 41-60.

Anderson, M.L. 2008. Multiple Inference and Gender Differences in the Effects of Early Intervention: A Re-evaluation of the Abecedarian, Perry Preschool, and Early Training Projects. *Journal of the American Statistical Association* 103: 1481-1495.

Andersson, C.I.M., C.G.K. Chege, E.J.O. Rao, and M. Qaim. 2015. Following Up on Smallholder Farmers and Supermarkets in Kenya. *American Journal of Agricultural Economics* 97: 1247-1266.

Angrist, J.D, and J.S. Pischke. 2009. *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton: Princeton University Press.

Angrist, J.D. 2006. Instrumental Variables Methods in Experimental Criminological Research: What, Why and How. *Journal of Experimental Criminology 2*(1): 23-44.

Angrist, J.D., G.W. Imbens, and D.B. Rubin. 1996. Identification of Causal Effects Using Instrumental Variables. *Journal of the American Statistical Association 91*(434): 444-455.

Asfaw, S., B. Shiferaw., F. Simtowe., and L. Lipper. 2012. Impact of Modern Agricultural Technologies on Smallholder Welfare: Evidence from Tanzania and Ethiopia. *Food Policy 37*(3): 283-295.

Ashraf, N., X. Giné, and D. Karlan. 2009. Finding Missing Markets (and a Disturbing Epilogue): Evidence from an Export Crop Adoption and Marketing Intervention in Kenya. *American Journal of Agricultural Economics 91*(4): 973-990.

Ayuya, O. I., E. O. Gido., H. K. Bett., J.K. Lagat., A.K. Kahi., and S. Bauer. 2015. Effect of Certified Organic Production Systems on Poverty Among Smallholder Farmers: Empirical Evidence from Kenya. *World Development* 67: 27-37.

Barrett, C.B. 2008. Smallholder Market Participation: Concepts and Evidence from Eastern and Southern Africa. *Food Policy 33*(4): 299-317.

Barrett, C.B., and M.R. Carter. 2010. The Power and Pitfalls of Experiments in Development Economics: Some Non-random Reflections. *Applied Economic Perspectives and Policy*, *32*(4): 515-548.

Bellemare, M.F., and L. Novak. 2017. Contract Farming and Food Security. *American Journal of Agricultural Economics 99*(2): 357-378.

Benjamini, Y., A.M. Krieger, and D. Yekutieli. 2006. Adaptive Linear Step-Up Procedures that Control the False Discovery Rate. *Biometrika* 93: 491-507.

Bloom, H.S. 2006. The Core Analytics of Randomized Experiment in Social Research. MDRC Working Paper. New York: MDRC

Bouis, H.E., and A. Saltzman. 2017. Improving Nutrition through Biofortification: A Review of Evidence from HarvestPlus, 2003 through 2016. *Global Food Security 12*: 49-58.

Buchinsky, M. 1998. Recent Advances in Quantile Regression Models: A Practical Guideline for Empirical Research. *Journal of Human Resources*: 88-126.

Carletto, C., A. Kirk, P.C. Winters, and B. Davis. 2010. Globalization and Smallholders: The Adoption, Diffusion, and Welfare Impact of Non-Traditional Export Crops in Guatemala. *World Development* 38: 814-827.

Carletto, C., P. Corral, and A. Guelfi. 2017. Agricultural Commercialization and Nutrition Revisited: Empirical Evidence from Three African Countries. *Food Policy* 67: 106-118.

Carletto, C., T. Kilic, and A. Kirk. 2011. Nontraditional Crops, Traditional Constraints: The Long-Term Welfare Impacts of Export Crop Adoption among Guatemalan Smallholders. *Agricultural Economics* 42: 61-75.

Carletto, G., M. Ruel, P. Winters., and A. Zezza. 2015. Farm-level Pathways to Improved Nutritional Status: Introduction to the Special Issue. *Journal of Development Studies* 51: 945-957.

Carter, M. R., R. Laajaj, and D. Yang. 2013. The Impact of Voucher Coupons on the Uptake of Fertilizer and Improved Seeds: Evidence from a Randomized Trial in Mozambique. *American Journal of Agricultural Economics 95*(5): 1345-1351.

Chege, C.G.K., C.I.M. Andersson, and M. Qaim. 2015. Impacts of Supermarkets on Farm Household Nutrition in Kenya. *World Development* 72: 394-407.

Chiputwa, B., and M. Qaim. 2016. Sustainability Standards, Gender, and Nutrition among Smallholder Farmers in Uganda. *Journal of Development Studies* 52: 1241-1257.

Coates, J., B. L. Rogers, A. Blau, J. Lauer, and A. Roba. 2017. Filling a Dietary Data Gap? Validation of the Adult Male Equivalent Method of Estimating Individual Nutrient

Intakes from Household-Level Data in Ethiopia and Bangladesh. *Food Policy* 72: 27-42.

Davis, K., E. Nkonya, E. Kato, D.A. Mekonnen, M. Odendo, R. Miiro, and J. Nkuba. 2012. Impact of Farmer Field Schools on Agricultural Productivity and Poverty in East Africa. *World Development 40*(2): 402-413.

de Brauw, A., P. Eozenou, and M. Moursi. 2015. Programme Participation Intensity and Children's Nutritional Status: Evidence from a Randomised Control Trial in Mozambique. *Journal of Development Studies 51*: 996-1015.

de Brauw, A., P. Eozenou, D. Gilligan, C. Hotz, N. Kumar, and J.V. Meenakshi. 2015. Biofortification, Crop Adoption and Health Information: Impact Pathways in Mozambique and Uganda. *HarvestPlus Working Paper* 21, Washington, DC: International Food Policy Research Institute.

de Groote, H., N.S. Gunaratna, M. Fisher, E.G. Kebebe, F. Mmbando, and D. Friesen. 2016. The Effectiveness of Extension Strategies for Increasing the Adoption of Biofortified Crops: The Case of Quality Protein Maize in East Africa. *Food Security 8*(6): 1101-1121.

de Haen, H., S. Klasen, and M. Qaim. 2011. What Do We Really Know? Metrics for Food Insecurity and Undernutrition. *Food Policy* 36: 760-769.

de Janvry, A., and E. Sadoulet. 2009. Agricultural Growth and Poverty Reduction: Additional Evidence. *The World Bank Research Observer 25*(1): 1-20.

Duflo, E., R. Glennerster, and M. Kremer. 2007. Using Randomization in Development Economics Research: A Toolkit. *Handbook of Development Economics* (Vol. 4, Chapter 61), 3895-3962.

Fanzo, J.C. 2017. Decisive Decisions on Production Compared with Market Strategies to Improve Diets in Rural Africa. *Journal of Nutrition 147*(1): 1-2.

FAO and FHI 360. 2016. *Minimum Dietary Diversity for Women: A Guide for Measurement*. Food and Agriculture Organization of the United Nations. Rome.

FAO, WHO, and UNU 2001. *Human Energy Requirements: Report of a Joint Expert Consultation*. Food and Nutrition Technical Report, Food and Agriculture Organization of the United Nations, Rome.

FAO. 2016. *Minimum Dietary Diversity for Women: A Guide for Measurement*. Food and Agriculture Organization of the United Nations, Rome.

FAO. 2017. *The State of Food Security and Nutrition in the World*. Rome: Food and Agriculture Organization of the United Nations.

Fischer, E., and M. Qaim. 2012a. Gender, Agricultural Commercialization, and Collective Action in Kenya. *Food Security* 4: 441-453.

Fischer, E., and M. Qaim. 2012b. Linking Smallholders to Markets: Determinants and Impacts of Farmer Collective Action in Kenya. *World Development 40*(6): 1255-1268.

Fisher, M., and V. Kandiwa. 2014. Can Agricultural Input Subsidies Reduce the Gender Gap in Modern Maize Adoption? Evidence from Malawi. *Food Policy 45*: 101-111.

Foster, A.D., and M.R. Rosenzweig. 2010. Microeconomics of Technology Adoption. *Annual Review of Economics 2*(1): 395-424.

Foster, J., J. Greer., and E. Thorbecke. 1984. A Class of Decomposable Poverty Measures. *Econometrica: Journal of the Econometric Society*: 761-766.

Gilligan, D.O. 2012. Biofortification, Agricultural Technology Adoption, and Nutrition Policy: Some Lessons and Emerging Challenges. *CESifo Economic Studies 58*(2): 405-421.

Goodhue, R.E., K. Klonsky, and S. Mohapatra. 2010. Can an Education Program be a Substitute for a Regulatory Program that Bans Pesticides? Evidence from a Panel Selection Model. *American Journal of Agricultural Economics 92*(4): 956-971.

Greene, W.H. 2012. *Econometric Analysis* (*7th Edition).* Prentice Hall: Upper Saddle River, NJ.

Guardabascio, B., and M. Ventura. 2014. Estimating the Dose–Response Function through a Generalized Linear Model Approach. *The Stata Journal* 14:141-158.

Haddad, L. J., J. Hoddinott, and H. Alderman. 1997. *Intrahousehold Resource Allocation in Developing Countries*. Johns Hopkins University Press, Baltimore, MD.

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

Hirano, K., and G.W. Imbens. 2004. The Propensity Score with Continuous Treatments. In *Applied Bayesian Modeling and Causal Inference from Incomplete-Data Perspectives: An Essential Journey with Donald Rubin's Statistical Family* (eds. A. Gelman and X.-L. Meng), John Wiley & Sons Ltd, Chichester, UK, 73-84.

Hoddinott, J., and L. Haddad. 1995. Does Female Income Share Influence Household Expenditures? Evidence from Cote d’Ivoire. *Oxford Bulletin of Economics and Statistics* 57: 77–96.

IFC. 2014. *Inclusive Business Case Study: Kenya Tea Development Agency Ltd (KTDA)*. International Finance Corporation.

IFPRI. 2017. *Global Nutrition Report*. Washington, DC: International Food Policy Research Institute.

Imbens, G.W., and J.M. Wooldridge. 2009. Recent Developments in the Econometrics of Program Evaluation. *Journal of Economic Literature* 47: 5–86.

IOM. 2006. *Dietary Reference Intakes: Applications in Dietary Assessment*. Institute of Medicine, National Academy Press, Washington, DC.

IZiNCG (International Zinc Nutrition Consultative Group). 2004. Assessment of the risk of zinc deficiency in populations and options for its control. *Food and Nutrition Bulletin*, 25, S91–S202.

Jones, A.D. 2017. On-Farm Crop Species Richness Is Associated with Household Diet Diversity and Quality in Subsistence-and Market-Oriented Farming Households in Malawi. *Journal of Nutrition* 147: 86-96.

Jones, A.D., A. Shrinivas., and R. Bezner-Kerr. 2014. Farm Production Diversity is Associated with Greater Household Dietary Diversity in Malawi: Findings from Nationally Representative Data. *Food Policy* 46: 1-12.

Jones, K.M., and A. de Brauw. 2015. Using Agriculture to Improve Child Health: Promoting Orange Sweet Potatoes Reduces Diarrhea. *World Development 74*: 15-24.

Kabunga, N.S., T. Dubois, and M. Qaim. 2012. Heterogeneous Information Exposure and Technology Adoption: The Case of Tissue Culture Bananas in Kenya. *Agricultural Economics 43*(5): 473-486.

Kassie, M., M. Jaleta, and A. Mattei. 2014. Evaluating the Impact of Improved Maize Varieties on Food Security in Rural Tanzania: Evidence from a Continuous Treatment Approach. *Food Security* 6: 217-230.

Kennedy, G., T. Ballard., and M.C. Dop. 2013. *Guidelines for Measuring Household and Individual Dietary Diversity*. Food and Agriculture Organization of the United Nations, Rome.

Key, N., E. Sadoulet, and A. de Janvry. 2000. Transactions Costs and Agricultural Household Supply Response. *American Journal of Agricultural Economics* 82: 245-259.

Kisii County Government. 2013. *The First County Integrated Development Plan 2013-2017 for Kisii County*. Kisii, Kenya.

Kisii County Government. 2013. *The First County Integrated Development Plan 2013-2017 for Kisii County*. Kisii, Kenya.

Kluve, J., Schneider, H., Uhlendorff, A., and Zhao, Z. 2012. Evaluating Continuous Training Programmes by Using the Generalized Propensity Score. *Journal of the Royal Statistical Society: Series A* 175 (2): 587-617.

KNBS. 2015. *Kenya Demographic and Health Survey 2014*. Kenya National Bureau of Statistics, Nairobi.

Koenker, R., and G. Bassett Jr. 1978. Regression Quantiles. *Econometrica: Journal of the Econometric Society*, 33-50.

Koenker, R., and K. F., Hallock. 2001. Quantile Regression. *Journal of Economic Perspectives 15*(4): 143-156.

Koppmair, S., M. Kassie, and M. Qaim. 2017. Farm Production, Market Access and Dietary Diversity in Malawi. *Public Health Nutrition* 20: 325-335.

Krishnan, P., and M. Patnam. 2013. Neighbors and Extension Agents in Ethiopia: Who Matters More for Technology Adoption? *American Journal of Agricultural Economics* 96: 308-327.

Lambrecht, I., B. Vanlauwe, R. Merckx, and M. Maertens. 2014. Understanding the Process of Agricultural Technology Adoption: Mineral Fertilizer in Eastern DR Congo. *World Development 59*: 132-146.

Läpple, D., and T. Hennessy. 2015. Assessing the Impact of Financial Incentives in Extension Programmes: Evidence from Ireland. *Journal of Agricultural Economics 66*: 781-795.

Lukmanji, Z., E. Hertzmark., N. Mlingi., V. Assey., G. Ndossi, and W. Fawzi, 2008. *Tanzania Food Composition Tables*. MUHAS-TFNC, HSPH, Dar es Salaam.

Magnan, N., D.J. Spielman, T.J. Lybbert, and K. Gulati. 2015. Leveling With Friends: Social Networks and Indian Farmers' Demand for a Technology with Heterogeneous Benefits. *Journal of Development Economics* 116: 223-251.

Masset, E., L. Haddad, A. Cornelius., and J. Isaza-Castro. 2012. Effectiveness of Agricultural Interventions that Aim to Improve Nutritional Status of Children: Systematic Review. *British Medical Journal* 344, d8222.

Mathenge, M. K., M. Smale., and J. Olwande. 2014. The Impacts of Hybrid Maize Seed on the Welfare of Farming Households in Kenya. *Food Policy 44*: 262-271.

Meemken, E.M., and M. Qaim (2018). Can Private Food Standards Promote Gender Equality in the Small Farm Sector? *Journal of Rural Studies* 58: 39-51.

Muricho, G., D. Manda., F. Sule., and M. Kassie. 2017. *Smallholder Agricultural Commercialization and Poverty: Empirical Evidence of Panel Data from Kenya.* Contributed Paper Prepared for Presentation at the 91th Annual Conference of the

Agricultural Economics Society, Royal Dublin Society in Dublin, Ireland 24 – 26 April 2017.

Muriithi, B.W., and J.A. Matz. 2015. Welfare Effects of Vegetable Commercialization: Evidence from Smallholder Producers in Kenya. *Food Policy* 50: 80-91.

Muthayya, S., J.H. Rah., J.D. Sugimoto., F.F. Roos., K. Kraemer., and R.E. Black. 2013. The Global Hidden Hunger Indices and Maps: An Advocacy Tool for Action. *PLoS ONE 8:* e67860.

Muyanga, M., and T.S. Jayne. 2008. Private Agricultural Extension System in Kenya: Practice and Policy Lessons. *Journal of Agricultural Education and Extension 14*(2): 111-124

Nyamira County Government. 2013. *The First County Integrated Development Plan 2013-2017 for Nyamira County*. Nyamira, Kenya.

Ogutu, S. O., T. Gödecke., and M. Qaim. 2017. *Agricultural Commercialization and Nutrition in Smallholder farm Households* (No. 97). GlobalFood Discussion Papers.

Olwande, J., M. Smale, M.K. Mathenge, F. Place, and D. Mithöfer. 2015. Agricultural Marketing by Smallholders in Kenya: A Comparison of Maize, Kale and Dairy. *Food Policy 52*: 22-32.

OPHI. 2017. *Oxford Poverty and Human Development Initiative 2017. Kenya Country Briefing*. Multidimensional Poverty Index Data Bank. OPHI, University of Oxford.

Pamuk, H., E. Bulte, and A.A. Adekunle. 2014. Do Decentralized Innovation Systems Promote Agricultural Technology Adoption? Experimental Evidence from Africa. *Food Policy 44*: 227-236.

Papke, L.E., and J.M. Wooldridge. 1996. Econometric Methods for Fractional Response Variables with an Application to 401 (k) Plan Participation Rates. *Journal of Applied Econometrics* 11:619-632.

Peterman, A., J.A. Behrman, and A. Quisumbing. R. 2014. A Review of Empirical Evidence on Gender Differences in Nonland Agricultural Inputs, Technology, and Services in Developing Countries. In *Gender in Agriculture* (pp. 145-186). Springer: Amsterdam.

Pingali, P., and N. Sunder. 2017. Transitioning Toward Nutrition-Sensitive Food Systems in Developing Countries. *Annual Review of Resource Economics 9*: 439-459.

Popkin, B.M., L.S. Adair., and S.W. Ng. 2012. Global Nutrition Transition and the Pandemic of Obesity in Developing Countries. *Nutrition Reviews* 70: 3-21.

Remans, R., F.A.J. DeClerck., G. Kennedy, and J. Fanzo. 2015. Expanding the View on the Production and Dietary Diversity Link: Scale, Function, and Change over Time. *Proceedings of the National Academy of Sciences USA* 112, E6082.

Rivers, D., and Q. H. Vuong. 1988. Limited Information Estimators and Exogeneity Tests for Simultaneous Probit Models. *Journal of Econometrics 39*(3): 347-366.

Sehmi, J. K. 1993. *National Food Composition Tables and the Planning of Satisfactory Diets in Kenya.* Kenya Government Press, Nairobi.

Sibhatu, K.T., and M. Qaim. 2017. Rural Food Security, Subsistence Agriculture, and Seasonality. *PLoS ONE 12*: e0186406.

Sibhatu, K.T., V.V. Krishna., and M. Qaim. 2015. Production Diversity and Dietary Diversity in Smallholder Farm Households. *Proceedings of the National Academy of Sciences USA* 112: 10657-10662.

Smith, L.C., and L. Haddad. 2015. Reducing Child Undernutrition: Past Drivers and Priorities for the Post-MDG Era. *World Development* 68: 180-204.

Smith, R. J.,and R. W. Blundell. 1986. An Exogeneity Test for a Simultaneous Equation Tobit Model with an Application to Labor Supply. *Econometrica: Journal of the Econometric Society*: 679-685.

StataCorp. 2013. *Stata: Release 13. Statistical Software.* College Station, TX: StataCorp.

Sununtnasuk, C., and J.L. Fiedler. 2017. Can Household-Based Food Consumption Surveys Be Used to Make Inferences about Nutrient Intakes and Inadequacies? A Bangladesh Case Study. *Food Policy* 72: 121-131.

Tipraqsa, P., and P. Schreinemachers. 2009. Agricultural Commercialization of Karen Hill Tribes in Northern Thailand. *Agricultural Economics* 40: 43-53.

Tschirley, D., T. Reardon, M. Dolislager, and J. Snyder. 2015. The Rise of a Middle Class in East and Southern Africa: Implications for Food System Transformation. *Journal of International Development* 27: 628-646.

United Nations. 2016. *The Sustainable Development Goals Report 2016*. United Nations, New York.

USAID. 2010. *Staple Foods Value Chain Analysis.* USAID Country Report-Kenya. http://www.fao.org/sustainable-food-value-chains/library/details/en/c/236730/

Verkaart, S., B.G. Munyua, K. Mausch, and J.D. Michler. 2017. Welfare Impacts of Improved Chickpea Adoption: A Pathway for Rural Development in Ethiopia? *Food Policy* 66: 50-61.

von Braun, J. 1995. Agricultural Commercialization: Impacts on Income and Nutrition and Implications for Policy. *Food Policy 20*(3): 187-202.

von Braun, J., and E. Kennedy., eds. 1994. *Agricultural Commercialization, Economic Development, and Nutrition*. Johns Hopkins University Press.

Wainaina, P., S. Tongruksawattana, and M. Qaim. 2016. Tradeoffs and Complementarities in the Adoption of Improved Seeds, Fertilizer, and Natural Resource Management Technologies in Kenya. *Agricultural Economics 47*(3): 351-362.

Waswa, L.M., I. Jordan, J. Herrmann, M.B. Krawinkel, and G.B. Keding. 2015. Community-Based Educational Intervention Improved the Diversity of Complementary Diets in Western Kenya: Results from a Randomized Controlled Trial. *Public Health Nutrition 18*: 3406-3419.

Wiesmann, U. M., B. Kiteme., and Z. Mwangi. 2014. *Socio-Economic Atlas of Kenya: Depicting the National Population Census by County and Sub-location.* Kenya National Bureau of Statistics, Centre for Training and Integrated Research in ASAL Development, Centre for Development and Environment.

Wooldridge, J. M. 2015. Control Function Methods in Applied Econometrics. *Journal of Human Resources 50(2)*: 420-445.

Wooldridge, J.M. 2002. *Econometric Analysis of Cross Section and Panel Data*. The MIT Press

World Bank. 2007. *World Development Report 2008: Agriculture for Development*. World Bank, Washington, DC.

World Bank. 2015. *Ending Poverty and Hunger by 2030: An Agenda for the Global Food System.* World Bank, Washington, DC.

World Bank. 2016. *Taking on Inequality: Poverty and Shared Prosperity 2016.* World Bank, Washington, DC.

World Bank. 2017. *World Development Indicators 2017*. Washington, DC. World Bank. https://openknowledge.worldbank.org/handle/10986/26447 (accessed 29.01.2018).

Zezza, A., C. Carletto, J.L. Fiedler, P. Gennari, and D. Jolliffe. 2017. Food Counts. Measuring Food Consumption and Expenditures in Household Consumption and Expenditure Surveys (HCES). Introduction to the Special Issue. *Food Policy* 72: 1-6.

Georg-August-Universität Göttingen

Africa Harvest BIOTECH FOUNDATION INTERNATIONAL

UNIVERSITY OF NAIROBI

Appendix B1. Survey Questionnaire

AGRICULTURE AND DIETARY DIVERSITY IN AFRICA: AN APPLICATION OF RANDOMISED CONTROLLED TRIALS IN KISII AND NYAMIRA, KENYA.

Goettingen University-Germany, University of Nairobi-Kenya and Africa Harvest Biotech Foundation International (Africa Harvest) are carrying out a research on different aspects of agricultural development. We are currently doing a survey which aims to provide more understanding about farmers' production and marketing decisions, and nutrition and health status. Your participation in answering these questions is very much appreciated. Your responses will be **COMPLETELY CONFIDENTIAL** and will only be used for research purpose. If you indicate your voluntary consent by participating in this interview, may we begin?

MODULE 0 – HOUSEHOLD ID

1	Household ID		8	County		12	**First visit date**	
2	Group ID		9	Sub-County			Result: 1=Interview completed 2= Interview partly completed 3= Specify	
3	Date of interview		9	Ward		13	**Second visit date**	
4	Start Time (24 Hr)		10	Village			Result: 1=Interview completed 2= Interview partly completed 3= Specify	
5	End time (24 Hr)		11	GPS Coordinates				
6	HH head Full Name							
7	Cell phone number							

Part 1: General Part (p.) Target Person: HH head or Group Member			**Part 2: Social Network (p.) Target Person: Group Member**				**Part 3: HH Food Consumption (p) Target Person: Women Responsible for HH Food**	
14	Respondent Name		21	Respondent Name		28	Respondent Name	
15	Respondent MEMID		22	Respondent MEMID		29	Respondent MEMID	
16	Enumerator Name		23	Enumerator Name		30	Enumerator Name	
17	Supervisor Name		24	Supervisor Name		31	Supervisor Name	
18	Date of interview		25	Date of interview		32	Date of interview	
19	Start Time (24 Hr)		26	Start Time (24 Hr)		33	Start Time (24 Hr)	
20	End time (24 Hr)		27	End time (24 Hr)		34	End time (24 Hr)	

Part 4: Childs questionnaire (p) Target Person: Mother of Child/Caretaker			**Part 5: Individual 1 (p) → First Day Target Person: Mother of Child**			**Part 6: Individual 2 (p) → First Day Target Person: Group Member**		
35	Respondent Name		44	Respondent Name		56	Respondent Name	
36	Respondent MEMID		45	Respondent MEMID		57	Respondent MEMID	
37	Child Name		46	Enumerator Name		58	Enumerator Name	
38	Child MEMID		47	Supervisor Name		59	Supervisor Name	
39	Enumerator Name		48	Date of interview		60	Date of interview	
40	Supervisor Name		49	Start Time (24 Hr)		61	Start Time (24 Hr)	
41	Date of interview		50	End time (24 Hr)		62	End time (24 Hr)	
42	Start Time (24 Hr)							
43	End time (24 Hr)		**Part 5: Individual 1 (p) → Second Day**			**Part 6: Individual 2 (p) →Second Day**		
Note Section:			51	Enumerator Name		63	Enumerator Name	
			52	Supervisor Name		64	Supervisor Name	
			53	Date of interview		65	Date of interview	
			54	Start Time (24 Hr)		66	Start Time (24 Hr)	
			55	End time (24 Hr)		67	End time (24 Hr)	

Questionnaire number (HHID) ____________________

We are researchers from University of Goettingen-Germany, University of Nairobi-Kenya, and Africa Harvest Biotech Foundation International (Africa Harvest). We are conducting research that aims to improve the knowledge on agriculture-nutrition linkages in the African small farm sector. We are particularly interested in understanding the mechanisms through which farmers can effectively adopt agricultural technologies that may improve their nutrition and health. We are currently conducting the first round of the survey and will do a follow-up about this time next year.

This informed consent is for smallholder farmers [like you] who belong to farmer groups and have engaged in farming activities during the last one year (October, 2014 to September, 2015). We are inviting you to participate in this research that mainly focuses on nutrition and health status of smallholder farmers in this area. We will ask you and some members of your household detailed questions on various topics related to agriculture, social networks, nutrition and health. We will also need to take measurements of the height and weight of selected adults and children below 5 years of age in your household. Your participation in this interview is entirely voluntary. Your responses will be treated with utmost confidentiality and the data will be used for research purposes only.

Do you have any questions that we need to clarify? [Make clarifications in case there are questions]If *No*, do you agree to take part in this survey, including the interviews and the measurements of adults and children?

If *Yes* let the potential respondent write name and sign below

Name____________________________________

Signature_________________________________

Questionnaire number (HHID) ____________________

TARGET PERSON: GROUP MEMBER OR HOUSEHOLD HEAD

MODULE 1: HOUSEHOLD DEMOGRAPHIC INFORMATION *(reference period between 1st Oct 2014 and 30th Sep 2015)*

Household composition***:*** *Please list all household members (All those who are under the care of household head in terms of food and shelter provision, and those who normally live and eat their meals together), starting with the household head.*

1	2	3	4	5	6	7	8	9	10	11	12
MEMID	**Name of the HH member**	**Gender** ***M = 1 F= 0***	**R/ship with HH head (*Codes A)***	**Age in years**	**Years of formal education (*Highest level attained*)**	**Marital Status *(Codes B)***	**Religion *(Codes C)***	**From which ethnic group are you? *(Codes D)***	**# of months in the last 12 months [NAME] has been away from home**	**Main Occupation based on time spent *(Codes E)***	**Household farm labour contribution *(for those above 16 years of age in the upper category) (Codes F)***
1											
2											
3											
4											
5											
6											
7											
8											
9											
10											
	Any other person born in this household and not listed above										
11											
12											
13											
14											

Code A
1= Head
2=Spouse
3=Son/daughter
4=Father/mother
5=Sister/brother
6=Grandchildren
7=Grandparents
8=Step children
9=Step parents
10 = Father/mother-in-law
11 =Sister/brother-in-law
12 = House girl
13 =Farm labourer
14 = Other relative
15= Other Unrelated

Code B
1= Married-monogamous
2= Married polygamous
3= Single
4= Divorced/separated
5= Widow/widower

Code C
1=Catholic
2=Protestant
3=SDA
4=Muslim
5=Traditionalist
6=No religion
77=Others (specify)

Code D
1=Kisii
77= Other (Specify)______

Code E
0= None
1= Farming (crop + livestock)
2= Casual labour on-other farm
3= Casual labour off-farm
4= Self-employed off-farm
5= Salaried employment (civil servant etc)
6=Student/school
77= Other (Specify)______

Code F
1= Part time
2= Fulltime
3=Does not work on farm

Questionnaire number (HHID) ___________________

MODULE 2: LAND HOLDING IN ACRES (*period between 1st Oct 2014 and 30th Sep 2015*)

2.1 How much land do you own in acres? ________________

2.2. How much of your total land is under homestead? _______________

2.3. Do you have a title deed for your land? ___________Yes=1 (all land), No=0 (no land), 3=partly

Land category	**Short rain season (Oct-Nov 2014)**		**Long rain season (Mar-Apr 2015)**	
	Cultivated	Fallow	Cultivated	Fallow
1. Own land (A)				
2. Rented in (B)				
3. Rented out (C)				
4. Total irrigated land				

2.4 What is the average cost of renting land per acre (Ksh/per year)? _________

CODES FOR MODULE 3

Codes A

1 Maize
2 Rice
3 Sorghum
4 Millet
5 Cassava
6 Beans (KKI5)
7 Field beans
8 Bananas
9 Cabbage
10 Cowpea
11 Groundnut
12 Soybean
13 Sweet potatoes
14 Orange Fleshed Sweet Potatoes (OFSP)
15 Black night shade
16 Sugarcane
17 Pineapple
18 Indigenous vegetables (specify____________
19 Indigenous vegetables (specify____________
20 Indigenous vegetables (specify_____________
21 Sukuma wiki (Kales)
22 Carrots
23 Passion Fruit
24 Irish potato
25 Bean leaves
26 Tea
27 Onion
28 Kales
29 Coffee
30 Napier grass
31 Avocado
77 Other__________________
78 Other__________________
79 Other___________________

Codes B

1. Improved
0. Local

Codes C

1. Kilogram
2. Litre
3. 90 Kg bag
4. 50 Kg bag
5. 25 Kg bag
6. Gorogoro
7. Debe
8. Wheelbarrow
9. Ox-cart
10. Bunch (bananas)
11. Piece/number
77 Other (specify)________

Questionnaire number (HHID) ___________________

MODULE 3: NON-LABOUR PURCHASED INPUT USE *(01.Oct 2014 to 30. Sept 2015 planting seasons, record separately by plots)*

1	2	3	4	5			6			7	8			9			10	
Plot Code (*Use alphabets in Cap*)	Crop Grown ***A***	Land under crop	Crop variety **B**	Seed **C**			Fertiliser(planting) (Fill once for intercrops) **C**			Oxen/ tractor hire cost	Farm manure (Fill once for intercrops ***C***			Pesticides/herbicides ***C***			Crop output ***C***	
				Qty	units	Price /Unit	Qty	Units	Price /Unit	Ksh	Qty	unit	Price /unit	Qty	Units	Price / unit	Qty	Unit
Short Rains																		
Long Rains																		
Perennial Crops																		

CODES FOR MODULE 4 (*period between 01. Oct 2014 and 30. Sept 2015)*

Codes A

1 Maize
2 Rice
3 Sorghum
4 Millet
5 Cassava
6 Beans (KKI5)
7 Field beans
8 Bananas
9 Cabbage
10 Cowpea
11 Groundnut
12 Soybean
13 Sweet potatoes
14 Orange Fleshed Sweet Potatoes (OFSP)
15 Black night shade
16 Sugarcane
17 Pineapple
18 Indigenous vegetables (specify____________
19 Indigenous vegetables (specify____________
20 Indigenous vegetables (specify_____________
21 Sukuma wiki (Kales)
22 Carrots
23 Passion Fruit
24 Irish potato
25 Bean leaves
26 Tea
27 Onion
28 Kales
29 Coffee
30 Napier grass
31 Avocado
77 Other__________________
78 Other__________________
79 Other__________________

Codes C

1. Kilogram
2. Litre
3. 90 Kg bag
4. 50 Kg bag
5. 25 Kg bag
6. Gorogoro
7. Debe
8. Wheelbarrow
9. Ox-cart
10. Bunch (bananas)
11. Piece/number
12. Not yet harvested (for perennials only)
77 Other (specify)________

Codes D

1. Farm gate
2. Village market
3. Main market
4. Collection center
77. Other (specify)________

Codes E

1. Own bicycle
2. Bodaboda
3. Hired truck
4. PSV
5. Donkey/oxen
6. Walking
7. Own truck
8. Taxi
77 Other (sp.)
99. NA

Code F: *1=male household head, 2= female household head, 3=female spouse, 4=joint decision, 5= others (specify)*___________

Questionnaire number (HHID) ___________________

MODULE.4: CROP UTILIZATION (*in the period between 01. Oct 2014 and 30 Sept 2015*)

1	2		3		4		5		6		7		8	9	10	11	12
Crop **Code A** (Aggregated crop)	Total Crop Output (Enter the total crop output from the plots)		Consumption		Saved as seed		Gift, tithe, donations, paid as wages		Quantity sold		Price		Point of most sales ***D***	Travel time to the point of sale (minutes)	Mode of transport **E**	Who mostly decides revenue use? **F**	Who mostly decides technology use e.g. variety **F**
	Qty	*Unit* ***C***	*Qty*	*Unit* ***C***	*Qty*	*Unit* ***C***	*Qty*	*Unit* ***C***	*Qty*	*Unit* ***C***	*Ksh*	*Unit* ***C***					
Short rain																	
Long rain																	
Perennial crop																	

Questionnaire number (HHID) ____________________

4.1 How easily can you access the market for sale of your produce (crop and or livestock)? *(Circe the applicable)*
1. Very easy 2. Easy 3.Difficult 4. Very difficult

4.2 Rank three most important market access constraints, if there exists any (***Prompt Codes G below***) 1.________________2.________________3.________________

Codes G:1. Poor infrastructure 2. Distant markets 3. Poor market prices 4. Cheating on quality standards/weighing scales 5. Lack of contracts or reliable buyers 6.Exploitative middlemen 77. Other (specify)________

MODULE 5: LABOUT INPUTS (*01.Oct 2014 to 30. Sept 2015 planting seasons, record hours worked by plot*)

1	2	3	4	5
Plot code	Plot size in acres	Plot manager *(F=0 M=1 Joint=3)*	Ploughing & harrowing Planting & thinning Applying fertiliser, Pesticide application (1st and 2nd) Weeding (1st and 2nd) Harvesting /Threshing/shelling/bagging	
			Family	Hired
Short Rains				
A				
B				
C				
D				
E				
F				
G				
H				
Long Rains				
A				
B				
C				
D				
E				
F				
G				
H				

5.6. What is the average <u>daily wage rate</u> for men and women in this village? Men________ Ksh/day Women__________Ksh/per day

MODULE 6: VARIETY/BREED AWARENESS AND UP-TAKE

	1	2	3	4	5	6	7	8	9	10	11	12	13
	New breed/variety/technologies	Have you ever heard of this variety/breed? *(1=Yes; 0=No) If No skip to the next technology*	Main source of information on the new variety/breed? **Codes A**	Have you ever planted /kept this variety/breed? *(1=Yes; 0=No) If YES, skip to 6*	If No to Q4, what was the main reason? **Codes C** *Then Skip to Q10*	What was the **main** source of breed kept/variety planted that year? **Codes B**	Number of seasons the variety has been planted, **since first** planting?	Number of years /variety/breed has been planted/kept	If you did not plant this variety/keep breed in 2015 what was the main reason? **Codes C**	Will you plant the variety/ keep the breed in future? *(1=Yes; 0=No, 88=don't know) if Yes skip to Q12*	What is the main reason? **Codes C**	Are you aware of the nutritional value of this variety or breed? (yes = 1, No = 0)	If yes to Q13 what was the source of information? **Code A**
1	*OFSP*												
2	*Butternut*												
3	*Kuroiler chicken*												
4	*Beans(KK15)*												

Code A	Code B	Code C	
1= Farmer Coop/Union 2= Farmer group 3=Extension staff/office 4= Other farmers (neighbours/relative) 5= Market (e.g. Agro vet/stockist) 6=Radio programs 7=Research centre (trials/demos) (*name* _____ 8= NGO/CBO *(name* ______ *9= Health centre/Practitioner* 77= Other*(specify* ______*)*	1= NGO free *(name ________)* 2= NGO subsidy *(specify_________)* 3= Extension staff demo plots 4= Other farmers 5= Market (Agrovet/local trader/stockist) 6=Farmer group/coop 7=Agricultural association/training centre 77= Other*(specify __________)*	1= Seed not available 2=Day old chicks not available 3=Lacked cash to buy seed/DOCs 4= Lacked credit to buy seed/DOCs 5= Prefer other varieties/breeds 6=Susceptible to diseases/pests 7=Poor taste 8=Low yielding/lays fewer eggs	9=Late maturing /longer maturity period 10=Low market prices/demand 11=High input requirements 12=Limited land to experiment/plant 13= limited information 77= Other*(specify _______)*

MODULE 7: VARIETY/BREED ATTRIBUTES, KNOWLEDGE & PERCEPTION

Instructions: Only ask the following questions to farmers who have ever heard or grown or kept the new technologies (listed below).

If *Yes*, ask for his/her perception of the performance of the technology (ies) against the listed attributes compared to his/her preferred local variety /breed. *Please mark the respondent's response with a* ***tick*** *in the appropriate cells below. If No skip to the next module.*

	1	**2**				**3**				**4**			
		OFSP				*Kuroiler chicken*				*Beans (KK15)*			
	Do you know the attributes of the following technologies? *Yes=1 No=0*	_________*If No Skip to the next technology, IF Yes ask for the attributes*				_________*If No Skip to the next technology, IF Yes ask for the attributes*				_________*If No Skip to the next technology, IF Yes ask for the attributes*			
	Technology attributes	*Better*	*Worse*	*No difference*	*Don't know*	*Better*	*Worse*	*No difference*	*Don't know*	*Better*	*Worse*	*No difference*	*Don't know*
1	Early maturity												
2	Yield												
3	Pest and disease resistance												
4	Marketability (demand)												
5	Cost of planting materials												
6	Market price received												
7	Cost of day old chicks												
8	Taste												
9	Lays more eggs												
10	Dry matter content												

7.5. If you have grown butternut before, what attributes did you consider? List three (*from the attributes above*), starting with the most important 1.____________2.________3.___________

7.6. Generally, which is your most important source of agricultural information 1 __________ *(Code A page 11)*

7.7. How easily can you obtain information on agricultural innovations from your most important source when needed? *(Circle the applicable)*

1. Very easy 2. Easy 3.Difficult 4. Very difficult

Questionnaire number (HHID) ____________________

MODULE 8: LIVESTOCK PRODUCTION AND MARKETING

8.1 For the last **12 months** ***(01. Oct 2014 to 30. Sep 2015)***, please give details of revenue and cost of livestock production?

(Please include all animals on the farm last year also those that were later sold or died) If no livestock is owned skip to next module)

	1	2a	2b	3a	3b	4a	4b	5a	5b	6	7	8	9	10	11	12
	Animal species	Stock at the beginning of the period (01.Oct.2014) (*If 0, skip to the next*)		Changes over the years: Home consumption		Sales		Stock at the end of 30.Sep.2015		Cash expenditures between 10/14 and 9/15Value in Ksh				Who decides sale?	Who decides revenue use?	Who decides technology use e.g. breed
		Unit	**Ksh**	**Units**	**Ksh**	**Units**	**Ksh**	**Units**	**Ksh**	**Veterinary treatment**	**Feed**	**Hired labor**	**Others**	**B**	**B**	**B**
1	Dairy cows/calves															
2	Cow/calves															
3	Goat															
4	Sheep															
5	Kuroiler/chicks															
6	Other chicken/chicks															
7	Donkeys															
8	Pigs															
9	Rabbits															
10	Ducks															
77	Other specify															
78	Other, specify															

8.2 For the last **12 months** ***(01. Oct 2014 to 30. Sep 2015)***, please give details of production and revenue of the following livestock products?

1	2		3		4		5	6	7
Animal product/services	Quantity produced		Quantity sold		Quantity Consumed		Price per unit	Who decides sale?	Who decides revenue use?
	Qty	Unit **A**	Qty	Unit **A**	Qty	Unit **A**		**B**	**B**
Milk									
Eggs									
Manure									
Honey									
Hide									
Others specify_________									

Code A: *1=litres, 2=Units/numbers, 3=Tray, 4=Kilogram, 5=50 kg bag , 6=90 kg bag, 7= Wheelbarrow, 8=pint (1/2 litre), 77=Other (specify)_______*

Code B: *1=male household head, 2= female household head, 3=female spouse, 4=joint decision, 77= others (specify)* ___________

Questionnaire number (HHID) ___________________

MODULE 9: HOUSEHOLD ASSETS (*Prompt for each item as listed below*)

9.1 As at September 2015, how many of the following items did the household own that are in **usable/repairable** condition?

To estimate the value ask the respondent how much they would be willing to buy the item in its current state if it were being sold to them

	ASSET	Total Quantity	Estimate total current value of the asset(s) if you were to buy it in its current state		ASSET	Total Quantity	Estimate total current value of the asset(s) if you were to buy it in its current state
1.	Tractor			**2.**	Slasher		
3.	Car/Van			**4.**	Axe		
5.	Pickup			**6.**	Panga		
7.	Motorcycle			**8.**	Hoes/Jembes		
9.	Bicycle			**10.**	Spades/shovel		
11.	Television			**12.**	Chemical spray pump		
13.	Radio			**14.**	Treadle pump		
15.	Mobile Phone			**16.**	Powered water pump		
17.	Refrigerator			**18.**	Mosquito net		
19.	Solar panels			**20.**	Greenhouse		
21.	Generator			**22.**	Water tank		
23.	Chaff cutter			**24.**	Store for farm produce		
25.	Ploughs for tractor			**26.**	Lanterns		
27.	Reaper			**28.**	Main house		
29.	ox-plough			**30.**	Wheelbarrow		
31.	Cart			**32.**	Computer/laptop		
33.	Livestock Kraal			**34.**	Biogas digesters		
35.	Other(specify_______			**36.**	Other(specify_______		
37.				**38.**			

Questionnaire number (HHID) ____________________

MODULE 10: CHARACTERISTICS OF MAIN HOUSE *(Instructions: please observe and ask about the following)*

1	2	3	4	5	6	7	8	9	10
Roofing material *1=grass /makuti 2=iron sheet 3=tiles 4=other, specify*	Wall material *1=mud 2=iron sheet 3=wood 4=plastered 5=bricks 6=stones 7=other specify*	Floor material *1= earth 2=cement 3=wood 4=tiles 5=other specify*	mode of ownership *1= owned by relative 2= rented 3= owned 4=other specify*	Type of toilet *1= bush 2= pit latrine 3= flush toilet 4= other specify*	Main source of water ***(Code A)***	Distance of the main source of water from the main house *in minutes by foot*	Mode of treating drinking water *1 = do nothing 2=Boil it 3=Use water guard/filter/ tablets) 4= Others (Specify*	Main cooking fuel *1=firewood 2=charcoal 3=paraffin 4=gas/biogas 5=electricity 6=solar power 7=other specify*	Main source of lighting *1=tin lamp 2=lantern 3=pressure lamp 4=electricity 5=solar power 6=other, specify*

Code A

1= piped into compound	*6=protected spring*	*11= roof catchments*
2= piped outside compound	*7= Borehole protected (private)*	*12= water hawkers-cart /boda boda*
3= stream/river	*8=borehole unprotected (private)*	*13= water tankers*
4= well	*9=Borehole protected (shared)*	*77= other, specify*
5=unprotected spring	*10=Borehole unprotected (shared)*	

MODULE 11: OTHER SOURCES OF INCOME AND TRANSFER

11.1. Do you have any other off-farm employment? (Please prompt the codes to make sure nothing is forgotten.)

1	2	3	4	5a	5b
MEM ID	Type of Occupation **A**	Average Number of days worked per month 10/14 – 9/15	Average Number of months worked per year 10/14 – 9/15	Earning per unit Ksh	**B**

***Code A**: 1: Agricultural labour (casual+permanent)2 Casual labour (non-agricultural) 3 Salary (Permanent non-agricultural employment)*
***Code B**: 1=Day, 2=Month, 3=Year, 4=Lump sum, payment, 77=other specify*

11.2 Do you have any other sources of income? (*Please prompt the codes to make sure nothing is forgotten.)*

1	2	3	4
Categories	Code	Type of occupation	Amount /value received between Oct14/ Sept 15/ *for small businesses ask for profit (+) losses (-)*
Remittances/gifts/transfers/food aid	1		
Pension	2		
Small business	3	Brick making	
	4	Carpentry	
	5	Construction	
	6	Grain mill	
	7	Handicrafts	
	8	Beverage, local brew	
	9	Sales in shop, petty trade	
	10	Transport	
	77	Other, specify______	
Sales of forest products	11	Sale of wood and charcoal,	
	12	Sale of wild nuts/fruits	
Other agric. income	13	Sale of crop residues	
	14	Leasing out land	
	15	Renting out oxen for ploughing	
	16	Hiring out machinery services to other farmers	
	17	Dividends (T-bills, bonds, shares)	

MODULE 12: NON-FOOD EXPENDITURE

Consider the **<u>last year (Oct 14 - Sept 15)</u>** generally how much has your HH spent on the items listed in a typical year (see specification indicated for each item)?

		1	2
		Read out: Please exclude Business Expenditures *Enter 99 if respondent does not know.*	How much did your household spend on [ITEM/SERVICE] during the last year (Oct. 14 – Sept 15)?
			Value in Khs
Non-food	1	Rent (housing)	
	2	Personal care supplies	
	3	Clothes, shoes and bags, accessories	
	4	Detergent/washing powder	
	5	Electricity	
	6	Other non-food	
Transportation + communication	7	Fuel, maintenance, insurance, and tax for motorbike/car	
	8	Public transport	
	9	Airtime (incl. MPESA)	
	10	Other transportation, communication	
	11		
	12		
Education	13	School fees, books, Student's dress/uniform, Tuition and rental fee	
	14	Other cost of schooling	
	15		
	16		
Health	17	Medicine, doctor fees	
	18	Other health cost	
	19		
	20		
Social	21	Celebration and funeral cost	
	22	Recreation and entertainment	
	23	Contributions (eg. Church, groups)	
	24	Tobacco (incl. snuff and miraa)	
	25	Insurance (eg. Car, life, health)	
	26	Remittances transferred to other HH	
	27	Other social cost	
	28		
	29		

MODULE 13: INFORMATION ON CREDIT ACCESS

13.1 Could you obtain credit if you needed it for the purpose of operational agricultural expenses (e.g. buying fertilizer paying for labour etc.)? ____________________*1=Yes, 0=No*

13.2 During the last ***12 months (Oct14 to Sep15)***, have you or any other household member received any credit to buy inputs, or received inputs on credit?_________________*1=Yes, 0=No*

13.3 If yes 13.2, how much did you receive in Ksh? (___________________)(*Include the value of inputs if inputs are provided on credit*)

13.4 How much went into purchasing inputs? (_______________) (*Include the value of inputs if inputs are provided on credit*)

Questionnaire number (HHID) ____________________

MODULE 15: ACCESS TO SOCIOECONOMIC INFRASTRUCTURE

1	2	3	4	5
Social facilities	**Distance to the nearest (km)**	**Most frequently used means of transportation to the facility** *(Use codes A below)*	**Travel time with most frequently used means of transportation (in minutes)**	**One way cost to travel there (Ksh)**
1. Murram road				
2. Tarmac road				
3. Village market				
4. Main Agricultural input market				
5. Main agricultural product market				
6. Health centre				
7. Agric. extension agent				

Code A: Means of transport Codes

1=Bicycle; 2=Motorbike; 3=Car; 4=Walk; 77= Others, (specify) _______________

MODULE 16: SELF ASSESSMENT OF RISK

16.1 How would you describe yourself? Are you generally willing to take risks, or do you try to avoid taking risks?

Unaweza kusema nini kuhusu wewe mwenyewe? Je, uko tayari kufanya mambo ambayo hayana uhakikisho kwa ujumla, au je, wewe hujaribu kuepuka mambo ambayo hayana uhakikisho?

Tafadhali chagua nambari kati ya sufuri (0) na Kumi (10), ambapo thamani ya (0) maana yake ni " siku zote wewe hujaribu kuepuka mambo ambayo hayana uhakikisho " na thamani ya (10) maana yake ni " uko tayari kikamilifu kufanya mambo ambayo hayana uhakikisho

Please choose a number on the scale between 0 and 10, where the value 0 means "always trying to avoid risks" and the value 10 means "fully prepared to take risks". (please circle chosen number)

Always trying to avoid risks										Fully prepared to take risks
0	1	2	3	4	5	6	7	8	9	10

MODULE 17: SHOCKS EXPERIENCENCED BY THE HOUSEHOLD

	1	2	3	4
	Please answer the following questions accordingly	Did you experience [NAME OF SHOCK] in the last three years? *1=Yes, 0=No* *If No Skip to the next shock*	If yes, how many times has it occurred in the last three years	What was the intensity of the last shock to this household? 1=Severe 2= Moderate 3=Mild
	Climatic shocks			
1	Drought			
2	Floods			
3	Frosts			
4	Hailstorm			
	Biological shocks			
5	Pests or diseases that affected crops before harvest			
6	Pests or diseases that led to storage losses			
	Economic shocks			
7	Large increase in agricultural input prices			
8	Large decrease in agricultural output prices			
9	Large increase in food prices			
	Other shocks			
10	Loss of family member			
11	Job loss			
12	Acute illness			
77	Other (specify___________________________)			

Questionnaire number (HHID) ____________________

TARGET PERSON: PERSON RESPONSIBLE FOR FOOD PREPARATION

MODULE 20: HOUSEHOLD FOOD CONSUMPTION

Firstly, we would like to ask the following four questions before we continue to ask you about your household food consumption. *(Only one answer possible)*

1	Who is mainly responsible for the food preparation in the household?	**A**	
2	Who is mainly responsible for food purchase in the household?	**A**	
3	Who is the main decision maker on food expenditure in the household?	**A**	
4	Who is the main decision maker on non-food expenditure in the household?	**A**	

Code A

1	Respondent	4	Respondent + spouse	7	Whole family	10	Other relatives
2	Spouse	5	Grandparents	8	Daughter	11	Other (non relatives)
3	Brother/sister	6	Family members not living in the village	9	Son		

Now we would like to ask about food consumption in the past seven days. Indicate how much of the following food items your household consumed, the prices in Ksh and the source of its origin (This is for all food consumed in the household, by all the people listed on demographic table in Module 1. INCLUDE food prepared at home but eaten outside. EXCLUDE meals prepared outside the home).

Firstly, ask how many people were present in the last 7 days? Please note down the number of household members in the following table. Please differentiate between female, male and children, as well as household members and visiting members. Fill in NA if a food item was not consumed in the last 7 days.

5	6	7	8	9	10	11	12
Household members				Visiting members			
Adults		**Children**		**Adults**		**Children**	
Male	*Female*	*Male*	*Female*	*Male*	*Female*	*Male*	*Female*
				13a	13b	13c	13d
In addition, we would like to know the age of the visiting children *(Please record the child's age separately, child is define as anybody under 18this section)*							

Questionnaire number (HHID) ____________________

	14	15	16	17	17a	17b	17c	18
	Food Items consumed in the past 7 DAYS	**How much in total did your household consume during the last 7 days?**	**Unit of quantities consumed** *(Use codes above A)*	**Source (*record quantities)***				**Average price per purchased unit Ksh…**
				Own production	*Purchased*	*Gift*	*Other, specify*	
1	**Staple foods**							
2	Maize green							
3	Maize grain							
4	Maize flour							
5	Yam Tuber							
6	Yam flour							
7	Cassava Tuber							
8	Cassava flour							
9	Orange fleshed sweet potato							
10	Other sweet potato							
11	Irish potato							
12	Irish potato chips							
13	Arrowroots							
14	Sorghum grain							
15	Sorghum Flour							
16	Millet grain							
17	Millet flour							
18	Brown rice							
19	White rice							
20	Wheat grain							
21	Wheat flour brown							
22	Wheat flour white							
23	Cooking banana							
24	**Other staple foods**							
25								
26								
27								
28	**Vegetables**							
29	Amaranth leaves (Emboga)							
30	Black night shade (Rinagu)							
31	Butternut							
32	Cabbage							

	14	15	16	17	17a	17b	17c	18
	Food Items consumed in the past 7 DAYS	**How much in total did your household consume during the last 7 days?**	**Unit of quantities consumed** *(Use codes above A)*	**Source (*record quantities)***				**Average price per purchased unit Ksh…**
				Own production	*Purchased*	*Gift*	*Other, specify*	
33	Carrot							
34	Cow pea leaves							
35	Cucumber							
36	Eggplant							
37	Jute mallow (Omotere)							
38	Kales							
39	Mushrooms							
40	Okra							
41	Onion							
42	Pepper							
43	Pumpkin							
44	Pumpkin leaves (Risosa)							
45	Spider plant (Chinsaga)							
46	Spinach							
47	Stinging nettle (rise)							
48	Sweet potato leaves							
49	Tomato							
50	Vine spinache (Enerema)							
51	**Other vegetables**							
52								
53								
54								
55	**Nuts and Pulses**							
56	Beans dry							
57	Beans fresh							
58	Black beans							
59	Cashew nut							
60	Green grams							
61	Groundnut (boild)							
62	Groundnut (roasted)							
63	Lentils							

Code A	3 Teaspoon	6 Kg	9 10 kg bag	12 90 kg bag	15 Hand (Banana)	18 ½ kg tin	21 Bundles	24 Plate	77 Others Specify
1 Liter	4 Tablespoon	7 Gram	10 25 kg bag	13 Debe	16 Piece/Counts	19 1 kg tin	22 Handful	25 Pint (500 ml)	78 Others Specify
2 Milliliter	5 Serving spoon	8 5 kg bag	11 50 kg bag	14 Bunch	17 ¼ kg tin	20 Ties	23 Cup	26 Gorogoro	79 Others Specify

Questionnaire number (HHID) ___________________

	14	15	16	17	17a	17b	17c	18
	Food Items consumed in the past 7 DAYS	**How much in total did your household consume during the last 7 days?**	**Unit of quantities consumed** *(Use codes above A)*	**Source (*record quantities)***				**Average price per purchased unit Ksh…**
				Own production	*Purchased*	*Gift*	*Other, specify*	
64	Peas (incl cowpea (Egesare), pigeon peas, green peas)							
65	Sesame seeds							
66	Soya meat							
67	Soybean							
68	Soybean flour							
69	**Other pulses and nuts**							
70								
71								
72								
73	**Fruits**							
74	Apple							
75	Avocado							
76	Coconut							
77	Guava							
78	Melon							
79	Orange							
80	Passion fruit							
81	Physalis/goose berry							
82	Pineapple							
83	Ripe bananas							
84	Ripe mango							
85	Ripe pawpaw							
86	Sugar cane							
87	**Other fruits**							
88								
89								
90								
91	**Meat and animal Products**							
92	Beef sausage							
93	Bush meat (Game meat)							
94	Chicken							
95	Chicken sausage							

	14	15	16	17	17a	17b	17c	18
	Food Items consumed in the past 7 DAYS	**How much in total did your household consume during the last 7 days?**	**Unit of quantities consumed** *(Use codes above A)*	**Source (*record quantities)***				**Average price per purchased unit Ksh…**
				Own production	*Purchased*	*Gift*	*Other, specify*	
96	Cow meat							
97	Eggs (pieces)							
98	Fish							
99	Goat/ Sheep meat							
100	Liver (from any animal)							
101	Offal's (matumbo)							
102	Pork							
103	Sardine (dagaa)							
104	Termites							
105	Turkey (batamzinga)							
106	**Other meats**							
107								
108								
109								
110	**Dairy products**							
111	Cheese							
112	Ice cream							
113	Milk (cow/goat milk)							
114	Powdered milk							
115	Sour milk (mala)							
116	Yoghurt							
117	**Other dairy product**							
118								
119								
120								
121	**Beverages**							
122	Cocoa powder							
123	Coffee (powder)							
124	Drinking chocolate							
125	Milo powder							
126	Soya powder							
127	Tea (leaves)							
128	**Other beverages**							
129								

Code A																	
		3	Teaspoon	6	Kg	9	10 kg bag	12	90 kg bag	15	Hand (Banana)	18	½ kg tin	21	Bundles	24	Plate	77	Others Specify
1	Liter	4	Tablespoon	7	Gram	10	25 kg bag	13	Debe	16	Piece/Counts	19	1 kg tin	22	Handful	25	Pint (500 ml)	78	Others Specify
2	Milliliter	5	Serving spoon	8	5 kg bag	11	50 kg bag	14	Bunch	17	¼ kg tin	20	Ties	23	Cup	26	Gorogoro	79	Others Specify

	14	15	16	17	17a	17b	17c	18
	Food Items consumed in the past 7 DAYS	**How much in total did your household consume during the last 7 days?**	**Unit of quantities consumed** *(Use codes above A)*	**Source (*record quantities*)**				**Average price per purchased unit Ksh…**
				Own production	*Purchased*	*Gift*	*Other, specify*	
130								
131	**Drinks**							
132	Apple juice							
133	Bottled beer							
134	Local beer							
135	Orange juice							
136	Pineapple juice							
137	Other juice (concentrates)							
138	Soft drinks (coke/fanta/etc)							
139	Wine							
140	**Other drinks**							
141	Water							
142								
143	**Condiments and spices**							
144	Salt							
145	Curry							
146	Ginger (tangawizi)							
147	Ketchup, Tomato sauce							
148	Pepper							
149	**Other Condiments and spices**							
150								
151								
152								
153	**Snacks**							
154	Bread							
155	Biscuit/cookies							
156	Popcorn							
157	Cakes							
158	Mandazi							
159	**Other snacks**							
160								

	14	15	16	17	17a	17b	17c	18
	Food Items consumed in the past 7 DAYS	**How much in total did your household consume during the last 7 days?**	**Unit of quantities consumed** *(Use codes above A)*	**Source (*record quantities*)**				**Average price per purchased unit Ksh…**
				Own production	*Purchased*	*Gift*	*Other, specify*	
161								
162	**Sugar and sweets**							
163	Sugar							
164	Chocolate							
165	Honey							
166	Sweets							
167	**Other sugar and sweets**							
168								
169								
170	**Fat and Oil**							
171	Animal fat							
172	Butter							
173	Corn oil							
174	Groundnut oil							
175	Margarine							
176	Sunflower Oil							
177	Vegetable oil							
178	Vegetables Fat							
179	**Other oil**							
180								
181								

Code A																			
1	Liter	3	Teaspoon	6	Kg	9	10 kg bag	12	90 kg bag	15	Hand (Banana)	18	½ kg tin	21	Bundles	24	Plate	77	Others Specify
2	Milliliter	4	Tablespoon	7	Gram	10	25 kg bag	13	Debe	16	Piece/Counts	19	1 kg tin	22	Handful	25	Pint (500 ml)	78	Others Specify
		5	Serving spoon	8	5 kg bag	11	50 kg bag	14	Bunch	17	¼ kg tin	20	Ties	23	Cup	26	Gorogoro	79	Others Specify

Questionnaire number (HHID) ____________________

TARGET PERSON: Mother or caretaker of child between the age of six to 59 months

MODULE 21: CHILD QUESTIONNAIRE – ONLY ONE CHILD WILL BE CONSIDERED

Please only ask this section to households that have children **under the age of 5 years**. Check first and probe if you are not sure! Ask the following questions to the person who is mainly responsible for the child care, e.g. mother, father, grandmother etc.)

If **NO** Child under 5 years is living in the household, please go to the next Module 22. If more than one child is under the age of 5 years, please choose the child of the group member. If not applicable choose child where respondent is available.

1	2	3a	3b	3c	4	5	6
MEMID of child	Name	Date of Birth *(please check with birth card; don't know=88)*			Age in month	Health card present? (No=0, Yes=1)	Which relationship to Child do you have? *(Respondent of this section, insert code below)* **Code A**
		Day	Month	Year			

Code A

1	Father	6	Stepfather/stepmother
2	Mother	7	Cousin
3	Grandmother/grandfather	8	Remote relative
4	Sister/brother	9	Care taker
5	Aunt/uncle	77	Others, specify______

PART 1: CHILD ANTHROPOMETRIC

In this section we would like to take the height, weight and mid-upper arm circumferences of your child. We will do this twice by two different enumerators to make sure to get the right weight and height.

1	1a	2	2a	3	3a	3b
Weight (kg)		Height (cm)		MUAC (measure tape)		
Data	Data	Data	Data	Data	Data	Data

Questionnaire number (HHID) ____________________

PART 2: CHILD HEALTH SECTION

Please ask the parent/caretaker of the child the following questions.

1	Where did you give birth to this child?	**A**	
2	Was the child a single or a multiple birth? *Single = 0, Twin = 1, Triple=3, Other, specify=77*		
3	What is the child's birth order e.g. was child born 1st 2nd, 3rd, etc.?		

4	Please record here, if the child received the following vaccines: *(insert 0 if they did not receive,"1" if they did & have card, and a "2" if the mother says the child received but there is no card), don't know =88*						
	a	Measles (at 9 month)		c	BCG (against tuberculosis) – at birth		
	b	Polio (OPV – at birth, 6wk, 10 wk and 14 wk)		d	DPT/DTaP (diphtheria, pertussis e.g. whooping cough and tetanus) (at 6wk, 10 wk and 14 wk)		

5	Has **the selected** child suffered from any of the following illnesses/symptoms in the *last 14 days*? *NO, skip to question 6* Indicate for how long ____ days											
	a	Diarrhoea		d	Fever		g	Blindness		j	Mouth problems	
	b	Measles		e	Fatigue/lethargy		h	Skin diseases		k	Eye disease	
	c	Anaemia		f	Respiratory illness		i	Pneumonia		l	Others, specify	

6	In the last year (1st Oct 2014 to 30th Sep 2015) did the child receive any of the following nutrition supplements or medical treatments (pills, liquids or supplemented food)? If YES, indicate for how long) ____ days; if No=0, Don't Know=88					
	a	Vitamin A		c	Iodine	
	b	Iron		d	Zinc	

7	Has [considered child] ever been breastfed? No=0, Yes=1, Don't Know=88				
8	Was [considered child] given anything to drink or eat BEFORE the first breastfeed? No=0, Yes=1, Don't Know=88				
9	Is child still being **breastfed?** No=0, Yes=1, Don't Know=88				
10	If the child is not breastfed anymore, at what age (in month) was breastfeeding stopped? *(please record age in month)*				
11	At what age was [considered child] given **other food** apart from breast milk? *Please verify by asking other household members; please record age in month or if not sure = 88, or if child does not take food yet =99*				
12	What do you think is the recommended period of exclusive breastfeeding? *Define exclusive breastfeeding ,if unsure, fill in 88 –* **Record Number of month**				
13	Which health problems or disease do you think are associated with **not exclusively breastfeeding**? *(Maximum 3 options)*	**B**	a	b	c

Code A

1	Health facility	6	Outside of home (e.g. in the field)
2	Home	7	At another home (e.g. relative, neighbor)
3	En route to facility	8	Government hospital
4	Private hospital	77	Other, specify
5	Other health care facility		

Code B

			Diarrhoea
1	Death	7	Weakness
2	Low weight for height	8	Low immunity
3	Low height for age	9	Obesity
4	Low weight	77	Other, specify
5	Stomach ache	99	None
6	Delayed achievement of development milestones (smiling, grabbing, etc)		

14	Has **any** child in the suffered from any of the following illnesses/symptoms in the *last 14 days*? If yes, please record the number of children affected:											
	a	Diarrhoea		d	Fever		g	Blindness		j	Mouth problems	
	b	Measles		e	Fatigue/lethargy		h	Skin diseases		k	Eye disease	
	c	Anaemia		f	Respiratory illness		i	Pneumonia		l	Others, specify	

Questionnaire number (HHID) ____________________

PART 3 Child: Dietary Recall from YESTERDAY (24hr DR) → INDICATE ONLY FOOD ITEMS

Now I would like to ask you about everything (including snacks) that (name of child) ate yesterday during the day or night, whether at home or outside of the home. *Continue through the day, **until respondent indicates child went to sleep** until the next day. If respondent mentions a mixed dish like a porridge, relish or stew, ask about all ingredients that went into the dish, including added oil, sugar or condiments.*

1	Which day of the week does this record represents?		**A**	
2	Is this a typical day? YES =1, NO=0, please specify			
3		4	5	6
Food item		Quantity -all foods-	Source -all food-**B**	Food preparation -all foods-**C**
Breakfast				
Snack				
Lunch				
Snack				
Dinner				
Snack				

Code A
1 Monday
2 Tuesday
3 Wednesday
4 Thursday
5 Friday
6 Saturday
7 Sunday

Code B
1 Own production
2 Purchased
3 Gift
77 Other, Specify

Code C
1 Raw
2 Dried
3 Boiled
4 Steamed
5 Cooked
6 Fried
7 Roasted
77 Other, specify

Questionnaire number (HHID) ____________________

1. TARGET PERSON: MOTHER/CARETAKER OF CONSIDERED CHILD

MODULE 22/1- FIRST INDIVIDUAL QUESTIONNAIRE (1)

This section will be answered by one chosen individual of the household, e.g. the mother of the child or the father. Please indicate MEMID as in the household questionnaire. If there was no child in the household, please interview the wife of the group member.

This module will be asked at two different days. At the first visit one 24-hour dietary recall will be asked. On the second visit a second 24 hour dietary recall and nutrition/health knowledge questions will be asked. Additionally, we would like to measure the respondent (weight, height, hip and waist). Please make an appointment with the respondent to come back after at least two days. Please make sure that the same person on both days is answering the questionnaire.

1
MEMID

	PART 3: INDIVIDUAL 1 ANTHROPOMETRIC						
	In this section we would like to take the height, weight and mid-upper arm circumferences of you. We will do this twice by two different enumerators to make sure to get the right weight and height.						
1	2	2a	3	3a	4	5	6
Type of clothe (light=0, heavy=1)	Weight (kg)		Height (cm)		Waist	Hip	MUAC (measure tape)
	Data	Data	Data	Data	cm	cm	Data

Questionnaire number (HHID) ____________________

PART 1 Mother/wife: Dietary Recall for YESTERDAY (1) (24hr DR) → Record Day

Now we would like to know, which food and drinks and how much of it did you (alone <u>*not*</u> *your household) consumed yesterday? Please indicate which food item, procession and where did you get it from? Write down all the foods and drinks mentioned. When composite dishes are mentioned, ask for the list of ingredients. When the respondents has finished, probe for the meals and snacks not mentioned.*

1	Which day of the week does this record represents?	**A**	
2	Is this a typical day? YES =1, NO=0, please specify		

3	4	5	6
Food item	Quantity -all foods-	Source -all food-**B**	Food preparation -all foods-**C**
Breakfast			
Snack			
Lunch			
Snack			
Dinner			
Snack			

Code A
1 Monday
2 Tuesday
3 Wednesday
4 Thursday
5 Friday
6 Saturday
7 Sunday

Code B
1 Own production
2 Purchased
3 Gift
77 Other, Specify

Code C
1 Raw
2 Dried
3 Boiled
4 Steamed
5 Cooked
6 Fried
7 Roasted
77 Other, specify

Questionnaire number (HHID) ____________________

Second visit (Day 2) **PART 2 Mother/wife: Dietary Recall for YESTERDAY (2) (24hr DR) → Record Day__________**

This part is for the second visit in the village. Please make sure you have the same person in front of you to answer the following questions .Now we would like to know, which food and drinks and how much of it did you (alone not your household) consumed yesterday? Please indicate which food item, procession and where did you get it from? Write down all the foods and drinks mentioned. When composite dishes are mentioned, ask for the list of ingredients. When the respondents has finished, probe for the meals and snacks not mentioned.

1	Which day of the week does this record represents?	**A**	
2	Is this a typical day? YES=1, NO=0, please specify		

3	4	5	6
Food item	Quantity -all foods-	Source -all food-**B**	Food preparation -all foods-**C**
Breakfast			
Snack			
Lunch			
Snack			
Dinner			
Snack			

Code A

1	Monday	4	Thursday	7	Sunday
2	Tuesday	5	Friday		
3	Wednesday	6	Saturday		

Code B

1	Own production
2	Purchased
3	Gift
77	Other, Specify

Code C

1	Raw	4	Steamed	7	Roasted
2	Dried	5	Cooked	77	Other, specify
3	Boiled	6	Fried		

Questionnaire number (HHID) ____________________

PART 2 Mother/Wife: INDIVIDUAL KNOWLEDGE QUESTIONS - <u>second visit (Day 2)</u>

No.	Question	Code				
1	During the last 4 weeks how would you rate your health?	**A**				
2	Do you suffer from any of the following diseases __________________ **[Read out Code B]**	**B**				
3	Have you suffered from any diseases or symptoms during the last 30 days? *Allow up to three answers*	**C**	a	b	c	d
4	Can you name anything that helps prevent you and other family members from getting diarrhoea? Yes=1, No=0, If No, skip to question 5					
	If yes, let respondent specify and tick the ones that apply underneath *– do not read out the list, probe for further responses – more than one answer possible*					
	a Washing hands		c Exclusive breast feeding		e Other specify	
	b Use latrine/ bury faces		d Protect food & water supplies with cover			

No.	Question	Code			
5	Which diseases or problems are caused by deficiency in iron? *(allow up to3 answers- Do not prompt)*	**D**	a	b	**c**
6	Do you know which foods have ample amounts of iron? *(allow up to3 answers – Do not prompt)*	**E**	a	b	**c**
7	Which diseases or problems are caused by deficiency in vitamin A? *(allow up to3 answers – Do not prompt)*	**F**	a	b	**c**
8	Do you know which foods have ample amounts of vitamin A? *(allow up to3 answers – Do not prompt)*	**G**	a	b	**c**
10	Did [NAME] receive any nutritious supplements in the last year (Oct 14 – Sept 15)? *Multiple answer possible, allow up to 3 answers*	**H**	a	b	c
11	Who do you think is the most informed person within your household concerning nutrition information?	**I**			
12	Who do you think is the most informed person within your household concerning agricultural information?	**I**			

No.	Question					
13	Have you heard of the following macronutrients: *No=0, Yes=1; If no, please go to question 24*	a Protein		B Carbohydrates		c Fat
	If yes, to question 20, please name which food items can supply those macronutrients: *(allow up to three answers)*			a	b	c
14		Protein	**J**			
15		Carbohydrates	**K**			
16		Fat	**L**			

No.	Question	Code
17	Are you aware of any health problems that are associated with eating none or too little of fresh fruits and vegetables*? Yes=1, No=0, Don't know=88*	
19	During the last year (Oct. 2014 - Sept 2015), have you noticed or received information about healthy eating or healthy diets? *No=0, Yes=1 (If NO, skip to next Individual)*	
20	If yes, where did you find, see or get this information on healthy eating/diet?	**M**

Code A
1 very good
2 Good
3 not good/not poor
4 a little poor
5 very poor
88 Don't know

Code B
1 Asthma
3 Cardiovascular disease
4 Diabetes
5 Hypertension
99 None

Code C
1 Bad teeth (ache)
2 Cholera
3 Diarrhoea
4 Ear/throat problem
5 Eye problems
6 Fainting
7 Fever
8 Flu/Cold
9 Headache
10 Hepatitis
11 High cholesterol
12 Intestinal worms
13 Malaria
14 Measles
15 Pneumonia
16 Skin Problem
17 Stomach ache
18 Tetanus
19 Tiredness/Fatigue
20 Tuberculosis
21 Vomiting
22 Weakness
77 Other (specify)
99 None

Code D
1 Fatigue, tiredness
2 Low concentration
3 Weak immune system
4 Shortage of blood
5 Reduction in intelligence
6 Small stature
7 Soreness of the moth
8 Unusual quantity of hair loss
77 Other, specify
88 Don't know

Code E
1 Green leafy vegetables
2 Peas
3 Soybean/chick peas/ lentils
4 Meat
5 Fish
6 orange/red coloured fruits & vegetables (eg., OFSP, tomatoes)
7 Liver
77 Other, specify

Code F
1 leading to eye blindness
2 Night blindness
3 Measles
4 Diarrhoea
5 Worms
77 Other, specify

Code G
1 Green leafy vegetables
2 Orange vegetables & fruits
3 Milk & milk products
4 Palm Oil
5 Eggs
6 Offal's
77 Other, specify
88 Don't know

Code H
1 Iodine
2 Iron
3 Vitamin A
4 Vitamin C
5 Zinc
77 Other ,specify
99 None

Code I
1 Respondent
2 Brother
3 Sister
4 Spouse
5 Grandparents
6 Female Child
7 Male Child
77 Other, specify

Code J
1 **Beans**
2 Beef
3 Cheese
4 Chicken
5 Dairy products
6 Eggs
7 Fish
8 Milk
9 Sardines
10 Yoghurt
77 Other, specify
88 Don't know

Code K -
1 Chapatti
2 Bread
3 Cassava
4 Yam
5 Crips
6 Potato
7 Plantain
8 Rice
9 Ugali
77 Other, specify
88 Don't know

Code L
1 Butter
2 Groundnuts
3 Lard
4 Kimbo
5 Oil
6 Palm Oil
77 Other, specify

Code M
1 Radio
2 TV
3 NGO
4 Doctor
5 Nutrition education (specify)
6 Newspaper
7 Health centre
8 Internet
9 Friends/Relatives
10 School
11 Community health worker
12 Church
77 Other Specify

Questionnaire number (HHID) ____________________

1. **TARGET PERSON: GROUP MEMBER (HUSBAND INCASE THE MOTHER IS THE GROUP MEMBER)**

MODULE 22/1- SECOND INDIVIDUAL QUESTIONNAIRE (2)

This section will be answered by one chosen individual of the household, the group member. Please indicate MEMID as in the household questionnaire.

This module will be asked at two different days. At the first visit one 24-hour dietary recall will be asked. On the second visit a second 24 hour dietary recall and nutrition/health knowledge questions will be asked. Additionally, we would like to measure the respondent (weight, height, hip and waist). Please make an appointment with the respondent to come back after at least two days. Please make sure that the same person on both days is answering the questionnaire.

1
MEMID

	PART 3: INDIVIDUAL 1 ANTHROPOMETRIC						
	In this section we would like to take the height, weight and mid-upper arm circumferences of you. We will do this twice by two different enumerators to make sure to get the right weight and height.						
1	2	2a	3	3a	4	5	6
Type of clothe (light=0, heavy=1)	Weight (kg)		Height (cm)		Waist	Hip	MUAC (measure tape)
	Data	Data	Data	Data	cm	cm	Data

Questionnaire number (HHID) ____________________

PART 1 Group member: Dietary Recall for YESTERDAY (1) (24hr DR) → Record Day

Now we would like to know, which food and drinks and how much of it did you (alone <u>not</u> your household) consumed yesterday? Please indicate which food item, procession and where did you get it from? Write down all the foods and drinks mentioned. When composite dishes are mentioned, ask for the list of ingredients. When the respondents has finished, probe for the meals and snacks not mentioned.

1	Which day of the week does this record represents?	**A**	
2	Is this a typical day? YES=1, NO=0, please specify		

3	4	5	6
Food item	Quantity -all foods-	Source -all food-**B**	Food preparation -all foods-**C**
Breakfast			
Snack			
Lunch			
Snack			
Dinner			
Snack			

Code A
1 Monday
2 Tuesday
3 Wednesday
4 Thursday
5 Friday
6 Saturday
7 Sunday

Code B
1 Own production
2 Purchased
3 Gift
77 Other, Specify

Code C
1 Raw
2 Dried
3 Boiled
4 Steamed
5 Cooked
6 Fried
7 Roasted
77 Other, specify

Questionnaire number (HHID) ____________________

Second visit (Day 2) **PART 2 Group member: Dietary Recall for YESTERDAY (2) (24hr DR) → Record Day_______**

This part is for the second visit in the village. Please make sure you have the same person in front of you to answer the following questions .Now we would like to know, which food and drinks and how much of it did you (alone not your household) consumed yesterday? Please indicate which food item, procession and where did you get it from? Write down all the foods and drinks mentioned. When composite dishes are mentioned, ask for the list of ingredients. When the respondents has finished, probe for the meals and snacks not mentioned.

1 Which day of the week does this record represents?		**A**	
2 Is this a typical day? YES =1, NO=0, please specify			
3	4	5	6
Food item	Quantity -all foods-	Source -all food-**B**	Food preparation -all foods-**C**
Breakfast			
Snack			
Lunch			
Snack			
Dinner			
Snack			

Code A

1	Monday	4	Thursday	7	Sunday
2	Tuesday	5	Friday		
3	Wednesday	6	Saturday		

Code B

1	Own production
2	Purchased
3	Gift
77	Other, Specify

Code C

1	Raw	4	Steamed	7	Roasted
2	Dried	5	Cooked	77	Other, specify
3	Boiled	6	Fried		

PART 2 Group member: INDIVIDUAL KNOWLEDGE QUESTIONS - <u>second visit (Day 2)</u>

1	During the last 4 weeks how would you rate your health?	**A**				
2	Do you suffer from any of the following diseases __________________ **[Read out Code B]**	**B**				
3	Have you suffered from any diseases or symptoms during the last 30 days? *Allow up to three answers*	**C**	a	b	c	D

4	Can you name anything that helps prevent you and other family members from getting diarrhoea? *Yes=1, No=0*								
	If yes, let respondent specify and tick the ones that apply underneath – *do not read out the list, probe for further responses – more than one answer possible*								
	a	Washing hands		C	Exclusive breast feeding		e	Other specify	
	b	Use latrine/ bury faces		d	Protect food & water supplies with cover				

5	Which diseases or problems are caused by deficiency in iron? *(allow up to3 answers- Do not prompt)*	**D**	a		b		**c**	
6	Do you know which foods have ample amounts of iron? *(allow up to3 answers - Do not prompt)*	**E**	a		b		**c**	
7	Which diseases or problems are caused by deficiency in vitamin A? *(allow up to3 answers - Do not prompt)*	**F**	a		b		**c**	
8	Do you know which foods have ample amounts of vitamin A? *(allow up to3 answers - Do not prompt)*	**G**	a		b		**c**	
10	Did you receive any nutritious supplements in the last year (Oct 14 – Sept 15)? *Multiple answer possible, allow up to 3 answers*	**H**	a		b		C	
11	Who do you think is the most informed person within your household concerning nutrition information?	**I**						
12	Who do you think is the most informed person within your household concerning agricultural information?	**I**						

13	Have you heard of the following macronutrients: *No=0, Yes=1*; *If no, please go to question 24*	a	Protein		B	Carbohydrates		c	Fat
14	If yes, to question 20, please name which food items can supply those macronutrients: *(allow up to three answers)*			a		b		C	
		Protein	**J**						
15		Carbohydrates	**K**						
16		Fat	**L**						

17	Are you aware of any health problems that are associated with eating none or too little of fresh fruits and vegetables*? Yes=1, No=0, Don't know=88*		
18	During the last year (Oct. 2014 - Sept 2015), have you noticed or received information about healthy eating or healthy diets? *No=0, Yes=1 (If NO, skip to next Individual)*		
19	If yes, where did you find, see or get this information on healthy eating/diet?	**M**	

Code A
- 1 very good
- 2 Good
- 3 not good/not poor
- 4 a little poor
- 5 very poor
- 88 Don't know

Code B
- 1 Asthma
- 3 Cardiovascular disease
- 4 Diabetes
- 5 Hypertension
- 99 None

Code C
- 1 Bad teeth (ache)
- 2 Cholera
- 3 Diarrhoea
- 4 Ear/throat problem
- 5 Eye problems
- 6 Fainting
- 7 Fever
- 8 Flu/Cold
- 9 Headache
- 10 Hepatitis
- 11 High cholesterol
- 12 Intestinal worms
- 13 Malaria
- 14 Measles
- 15 Pneumonia
- 16 Skin Problem
- 17 Stomach ache
- 18 Tetanus
- 19 Tiredness/Fatigue
- 20 Tuberculosis
- 21 Vomiting
- 22 Weakness
- 77 Other (specify)
- 99 None

Code D
- 1 Fatigue, tiredness
- 2 Low concentration
- 3 Weak immune system
- 4 Shortage of blood
- 5 Reduction in intelligence
- 6 Small stature
- 7 Soreness of the moth
- 8 Unusual quantity of hair loss
- 77 Other, specify
- 88 Don't know

Code E
- 1 Green leafy vegetables
- 2 Peas
- 3 Soybean/chick peas/ lentils
- 4 Meat
- 5 Fish
- 6 orange/red coloured fruits & vegetables (eg., OFSP, tomatoes)
- 7 Liver
- 77 Other, specify

Code F
- 1 leading to eye blindness
- 2 Night blindness
- 3 Measles
- 4 Diarrhoea
- 5 Worms
- 77 Other, specify

Code G
- 1 Green leafy vegetables
- 2 Orange vegetables & fruits
- 3 Milk & milk products
- 4 Palm Oil
- 5 Eggs
- 6 Offal's
- 77 Other, specify
- 88 Don't know

Code H
- 1 Iodine
- 2 Iron
- 3 Vitamin A
- 4 Vitamin C
- 5 Zinc
- 77 Other ,specify
- 99 None

Code I
- 1 Respondent
- 2 Brother
- 3 Sister
- 4 Spouse
- 5 Grandparents
- 6 Female Child
- 7 Male Child
- 77 Other, specify

Code J
- 1 **Beans**
- 2 Beef
- 3 Cheese
- 4 Chicken
- 5 Dairy products
- 6 Eggs
- 7 Fish
- 8 Milk
- 9 Sardines
- 10 Yoghurt
- 77 Other, specify
- 88 Don't know

Code K -
- 1 Chapatti
- 2 Bread
- 3 Cassava
- 4 Yam
- 5 Crips
- 6 Potato
- 7 Plantain
- 8 Rice
- 9 Ugali
- 77 Other, specify
- 88 Don't know

Code L
- 1 Butter
- 2 Groundnuts
- 3 Lard
- 4 Kimbo
- 5 Oil
- 6 Palm Oil
- 77 Other, specify

Code M
- 1 Radio
- 2 TV
- 3 NGO
- 4 Doctor
- 5 Nutrition education (specify)
- 6 Newspaper
- 7 Health centre
- 8 Internet
- 9 Friends/Relatives
- 10 School
- 11 Community health worker
- 12 Church
- 77 Other Specify

THANK YOU VERY MUCH FOR YOUR TIME