Xavier Gine
Shreena Patel
Cristina Cuellar-Martinez
Sandi McCoy
Lauren Ralph

Enhancing food production and food security through improved inputs

An evaluation of Tanzania's National Agricultural Input Voucher Scheme with a focus on gender impacts

February 2015

Impact Evaluation Report 23

Agriculture

International Initiative for Impact Evaluation

About 3ie

The International Initiative for Impact Evaluation (3ie) is an international grant-making NGO promoting evidence-informed development policies and programmes. We are the global leader in funding and producing high-quality evidence of what works, how, why and at what cost. We believe that better and policy-relevant evidence will make development more effective and improve people's lives.

3ie impact evaluations

3ie-supported impact evaluations assess the difference a development intervention has made to social and economic outcomes. 3ie is committed to funding rigorous evaluations that include a theory-based design, use the most appropriate mix of methods to capture outcomes and are useful in complex development contexts.

About this report

3ie accepted the final version of this report, 'Enhancing food production and food security through improved inputs: an evaluation of Tanzania's National Agricultural Input Voucher Scheme with a focus on gender impacts', as partial fulfilment of requirements under grant OW2.171 issued under Open Window 2. The content has been copyedited and formatted for publication by 3ie. Due to unavoidable constraints at the time of publication, a few of the tables or figures may be less than optimal. All of the content is the sole responsibility of the authors and does not represent the opinions of 3ie, its donors or its board of commissioners. Any errors and omissions are also the sole responsibility of the authors. All affiliations of the authors listed in the title page are those that were in effect at the time the report was accepted. Any comments or queries should be directed to the corresponding author Sandi McCoy at smccoy@berkeley.edu.

Funding for this impact evaluation was provided by 3ie's donors, which include UKaid, the Bill & Melinda Gates Foundation, Hewlett Foundation and 12 other 3ie members that provide institutional support. A complete listing is provided on the 3ie website at http://www.3ieimpact.org/about-us/3ie-members.

Suggested citation: Gine, Xavier, Patel, Shreena, Cuellar-Martinez, Cristina, McCoy, Sandi and Ralph Lauren, Enhancing food production and food security through improved inputs: an evaluation of Tanzania's National Agricultural Input Voucher Scheme with a focus on gender impacts, *3ie Impact Evaluation Report 23.* New Delhi: International Initiative for Impact Evaluation (3ie).

3ie Impact Evaluation Report Series executive editors: Jyotsna Puri and Beryl Leach
Managing editor: Paromita Mukhopadhyay
Assistant managing editor: Kanika Jha
Production manager: Omita Goyal
Assistant production manager: Pradeep Singh
Copy editor: Payal Dhar
Proofreader: Mathew PJ
Cover design: John F McGill
Printer: VIA Interactive
Cover photo: Enviu

© International Initiative for Impact Evaluation (3ie), 2015

Enhancing food production and food security through improved inputs: an evaluation of Tanzania's National Agricultural Input Voucher Scheme with a focus on gender impacts

Xavier Gine
The World Bank

Shreena Patel
The World Bank

Cristina Cuellar-Martinez
The World Bank

Sandi McCoy
University of California, Berkeley School of Public Health

Lauren Ralph
University of California, Berkeley School of Public Health

February 2015
3ie Impact Evaluation Report 23

International Initiative for Impact Evaluation

Abstract

The National Agricultural Input Voucher Scheme (NAIVS), administered since 2008 by the Ministry of Agriculture, Food Security and Cooperatives (MAFC), provides a 50 per cent subsidy for the purchase of chemical fertiliser and improved seed to maize and rice farmers in high-potential areas of Tanzania. The programme aims to improve access to and adoption of critical agricultural inputs in order to boost food production and reduce pressure on prices of food staples, thereby increasing incomes and bolstering food security. The input package provided by NAIVS contains vouchers for chemical fertiliser (basal and top dressing) and improved maize or rice seed. Households are selected to become beneficiaries by a Village Voucher Committee (VVC) based on specific eligibility criteria.

A rigorous impact evaluation was designed to determine NAIVS's impact on household incomes, production and food security. In addition, differential effects of the programme by gender of the household head were explored within each of these outcomes. The evaluation design centred on a set of targeting interventions, which both introduced the random assignment of beneficiaries in a set of villages and allowed the study to test the targeting efficiency of NAIVS. The interventions consisted of a public meeting to identify eligible farming households and a lottery to randomly select beneficiaries from among the list of those eligible. Villages were divided into four treatment groups, where PL villages received both the public meeting and lottery, PV received just the public meeting, VL received just the lottery, and VV received neither. Primary findings were drawn from a two-year panel survey of households, half of which were new NAIVS beneficiaries in the 2010–2011 planting season. The analysis focuses on Meru district in the Arusha region, which was the only area where the targeting interventions were conducted as planned.

Overall, the evaluation found that, in VV villages, improved input use and yields were higher among beneficiaries than non-beneficiaries. The study did not find impacts on household welfare indicators, such as incomes, food security or educational attainment. However, this may be due to the fact that follow-up data was collected only one year after baseline and these types of benefits required more time to accrue. Similarly, the evaluation found no differential impacts on women across the outcomes examined. The study also found that the targeting criteria were not enforced consistently by the VVC, indicating that beneficiaries in VV villages did not represent the target group; approximately 44 per cent of the overall sample of 'eligible' households did not meet the criteria. In villages where the lottery was used to select beneficiaries, there was evidence of selling or sharing of voucher inputs. Accordingly, yields were higher among non-beneficiaries in PL and VL villages compared to VV villages, probably because beneficiaries selected by lottery were passing on improved inputs to non-beneficiaries.

These findings indicate that the set of farming households expressly targeted by NAIVS was either not willing or not able to make use of the voucher inputs, and that the VVC, when left to its own targeting practices, was able to select the beneficiaries who would make best use of the inputs. Thus, in order to enhance efficiency while providing for the most in need, the NAIVS eligibility criteria should continue to be strictly enforced, but while also allowing the transfer and sale of vouchers or its contents, something that the current programme implementers vehemently oppose. Assuming the VVC was not targeting farmers who had already been purchasing improved inputs commercially, thereby resulting in NAIVS simply displacing these purchases, this recommendation would balance goals of targeting accuracy and efficiency while achieving the intended outcomes of increased adoption of improved inputs and higher yields.

Contents

List of figures and tables

Abbreviations and acronyms

DAP	diammonium phosphate
FI	food insecurity
HDDS	Household Dietary Diversity Score
HFIAS	Household Food Insecurity Access Scale
LPM	linear probability model
MAFC	Ministry of Agriculture, Food Security and Cooperatives
MRP	Minjingu rock phosphate
NAIVS	National Agricultural Input Voucher Scheme
NMB	National Microfinance Bank
OLS	ordinary least squares (regression)
OPV	open pollinated varieties
REPOA	Research on Poverty Alleviation
TASAF	Tanzania's Social Action Fund
VEO	village executive officer
VVC	Village Voucher Committee

1. Introduction

1.1 Context and intervention

Agriculture is central to the Tanzanian economy and is inextricably linked to food security, household income, poverty reduction and health. It accounts for 27 per cent of GDP, 80 per cent of employment and 75 per cent of rural household incomes, and provides over 95 per cent of Tanzania's annual food requirement. Women's labour plays a fundamental role in Tanzania, 92 per cent of rural women are engaged in agriculture (United Republic of Tanzania 2008; World Bank 2009) and women produce up to 80 per cent of the continent's food (Rowling 2008). However, a persistent challenge for all farmers in Tanzania (and across Africa) is low yields; agriculture is almost entirely rain fed, crop yields are 20 to 30 per cent of potential, and improved inputs are used at extremely low rates despite their high expected rates of return (Duflo, Kremer and Robinson 2008; Suri, forthcoming). Further, despite their prominent role in agriculture, women farmers face additional challenges, including harmful gender norms, weak property and contractual rights, lack of initial investments for inputs, and overall unequal access to information and extension services (Quisumbing and Pandolfelli 2008; World Bank 2009).

In response to these issues, the World Bank's Accelerated Food Security Project (AFSP) recently scaled up Tanzania's National Agricultural Input Voucher Scheme (NAIVS) to 2.5 million households across the country. The NAIVS programme, which began in 2008, provides farmers growing rice and maize with vouchers for 50 per cent subsidies on fertilisers and improved seeds in order to boost food production, reduce pressure on food staple prices, increase incomes and bolster food security. While agricultural subsidies have received criticism in the past due to their high cost and risk of market distortions, the recurring instances of food shortages and rising food prices in developing countries has led to renewed interest in subsidies as a way to combat these problems, particularly in Africa (Duflo, Kremer and Robinson, forthcoming). Agricultural input subsidies are appealing because they encourage sustainable changes in farming practices that can result in long-term productivity increases to bolster food security and put downward pressure on food prices. The main barriers to the use of chemical fertiliser and improved seed varieties are credit constraints, risk aversion and lack of awareness, to which women are particularly vulnerable (Feder, Just and Zilberman 1985). Input subsidies address farmers' credit constraints directly and they also produce a demonstration effect of the benefits of improved input use, reducing other farmers' perception of risk and spreading knowledge to farmers who were previously unaware of these technologies.

1.2 Gender and agriculture

Given their central role in agriculture, supporting women farmers specifically has been widely recognised as a strategy to positively impact agricultural productivity, improve food security and decrease susceptibility to deleterious health effects for themselves and their families. Women's position as primary caregivers within most rural African households suggests that supporting women's agricultural productivity could have beneficial effects on their own nutritional and health outcomes as well as that of their children, including a reduction in vitamin deficiencies, stunted growth and severe malnutrition. The bidirectional nature of these effects is also compelling. When the health of a woman (or that of her children) is compromised, her ability to work and provide economic and other support for her family is similarly compromised. For example, the effects of high HIV prevalence on regional agricultural

productivity have been well documented (Gillespie and Drimie 2009). Thus, any strategy to enhance agricultural productivity and achieve sustainable changes in farming practices must also address the unique challenges and constraints faced by women, who provide a significant proportion of agricultural labour.

1.3 Evaluation of National Agricultural Input Voucher Scheme, with a focus on targeting and gender

A rigorous impact evaluation was designed to determine NAIVS's impact on household incomes, production, food security and self-reported health status. In addition, differential effects of the programme by the gender of the household head would be explored within each of these outcomes. The evaluation design centred on a set of targeting interventions, which would both introduce the random assignment of beneficiaries in a set of villages and allow the study to test the targeting efficiency of NAIVS. Primary evaluation findings are drawn from a two-year panel survey of households, half of which were new NAIVS beneficiaries in the 2010–2011 planting season.

The study findings have important implications for policymaking around input subsidies as a mechanism to enhance productivity, reduce poverty and improve food security on a large-scale and across a variety of agro-ecological zones. Evaluation results provide a critical and timely opportunity to assess the efficacy of input subsidies and provide lessons to other developing countries, especially in Africa, who are considering scaling up investments in the form of subsidies. While input subsides have been used and are being considered by many countries, there is very little rigorous evidence of their impact and, in Tanzania, the scheme already absorbs a significant portion of the budget of the Ministry of Agriculture, Food Security and Cooperatives (MAFC).

1.4 Evaluation questions

1) What is the impact of NAIVS on:
   a) Production and welfare at the beneficiary level?
      i. Direct returns: input use, productivity, profitability, quantity of output sold on the market.
      ii. Indirect returns: nutrition, household welfare, human capital accumulation.
   b) Heterogeneous treatment effects and spillovers within villages?
      i. Are returns to fertiliser and input use heterogeneous by gender of the household head?
      What were barriers and facilitators to NAIVS participation among female farmers?
      ii. How are the non-eligible or non-participating households affected by the programme?
2) How effective was the NAIVS programme targeting method?
   a) Do the village voucher committees target voucher beneficiaries according to the programme criteria?
   b) Does the introduction of participatory village meetings to identify eligible farmers and a lottery to select beneficiaries result in less efficient targeting?
3) What is the cost-effectiveness of the NAIVS programme?[1]

[1] A full cost-effectiveness analysis of the programme is currently being completed by researchers at the World Bank in Tanzania. The paper is available on the website of the Agriculture Public Expenditure programme administered by the World Bank.

1.5 Report overview

This report will analyse the results of the NAIVS impact evaluation, with a particular focus on the findings from the targeting intervention and their implications for the effectiveness of the programme. Chapter 2 provides a description of the NAIVS programme and targeting interventions being tested. Chapter 3 describes the evaluation design, including sampling and questionnaires. Chapter 4 provides a summary of programme implementation, while chapters 5 and 6 provide impact results and policy recommendations respectively.

2. Programme description

2.1 National Agricultural Input Voucher Scheme

The NAIVS programme is implemented by Tanzania's Ministry of Agriculture, Food Security and Cooperatives and was developed in response to the food crisis that peaked in the late 2000s and resulted in higher prices of major cereals, pulses and farm inputs. The goal of the programme is to improve access to and adoption of critical agricultural inputs in order to boost food production and reduce pressure on prices of food staples, thereby increasing incomes and bolstering food security. The key intervention of NAIVS is a subsidised package of improved inputs targeted to a subset of maize and rice farmers. The programme began in 2008 with 750,000 beneficiary farmers and has been scaled up each year since. In the 2010–2011 season, NAIVS reached 2 million farmers and, in 2011–2012, expanded to reach the full target of 2.5 million farmers in over 65 districts across the country.

2.1.1 Intervention

The NAIVS package provides vouchers for three improved agricultural inputs:
(1) basal phosphate fertiliser (diammonium phosphate, DAP, or Minjingu rock phosphate, MRP), (2) urea top dressing, and
(3) improved maize (hybrid or open pollinated variety) or paddy seed.
This package is designed for use on one acre of land. Households are selected to become beneficiaries by a Village Voucher Committee (VVC) that must be formed in each participating village. The VVC identifies eligible farming households and selects beneficiary households according to the following stipulated programme criteria:
(1) the farming household must not cultivate more than one hectare of maize or rice;
(2) the farming household must be able to afford the subsidised cost of the input package (also known as the 'top-up' amount);
(3) priority should be given to female-headed farming households; and
(4) priority should be given to farming households that have not used improved inputs in the previous five years.

Village Voucher Committees select beneficiaries prior to the start of the planting season based on the allocation of vouchers received that year, and beneficiaries should receive the subsidy for a three-year period. Farmers are expected to 'graduate' from the programme after three years, with the hope that they will continue to purchase inputs commercially. Selected beneficiaries receive three vouchers, each stating its face value and the specific input type that it is intended for. Each voucher has a face value of approximately half the cost of the input in the given region for the given year, as determined by MAFC before the start of the planting season, and can be redeemed with agro-dealers who have been certified to handle vouchers by the local district government. Farmers must present the voucher along with the additional top-

up payment to the agro-dealer in order to receive the inputs; agro-dealers then submit vouchers to the local National Microfinance Bank (NMB) branch in order to be reimbursed for the face value of the voucher.

2.1.2 Theory of change

The anticipated causal pathways between NAIVS and each of the agriculture, nutrition and welfare outcomes have been outlined in the attached Theory of Change (see Appendix A), which guided data collection for the evaluation. The introduction of the subsidised package of inputs is expected to result in increased adoption of improved seeds and chemical fertiliser for maize and rice cultivation, which will, in turn, produce higher maize and rice yields. On a macro level, higher agricultural productivity will result in a larger supply of food grains and allow food prices to fall. It will also support food security across Tanzania, assuming food grain surpluses in highly productive areas are transferred to food insecure parts of the country. On a micro level, increased agricultural productivity will improve household food security and translate to increased incomes from the sale of surplus harvest. Increased incomes are expected to facilitate the purchase of more nutritious foods, leading to a reduction in malnutrition. They may also be used for household expenses relating to health and education, thereby improving human capital accumulation. Increased incomes are key for the sustainability of the immediate project goal of increasing the adoption and use of improved inputs, as the subsidy only lasts for three years per household and farmers have to purchase inputs commercially thereafter. Finally, increased incomes are expected to make households more economically stable and less vulnerable to shocks, which is expected to result in a reduction in poverty.

The theory of change rests upon the following key assumptions:

- Targeting criteria are upheld and the beneficiary group consists of farmers who were not previously using improved inputs, otherwise the subsidy may displace commercial purchases that were already taking place;
- Beneficiaries understand and trust the technology enough to use the vouchers to purchase the intended inputs;
- Beneficiaries have sufficient skill and knowledge to use the inputs correctly; and
- Weather conditions are suitable for cultivation and harvest.

2.1.3 Key indicators

The following table outlines the key indicators included in the evaluation to track expected outcomes and critical assumptions:

Table 1: Key indicators in the evaluation

Level	Result	Indicator
Output	Vouchers distributed	• Voucher use
Outcomes	Increased adoption of improved inputs Improved yields, increased production Improved household welfare	• Plot-level input use (basal fertiliser, urea and seed) • Maize and rice yields • Agricultural income • Household consumption • Food security index • Dietary diversity index • School attendance/enrolment • Household expenditures
Assumptions	Vouchers do not displace commercial purchases of inputs Farmers know how to use inputs	• Previous input use • Knowledge of fertiliser application practices

2.2 Targeting interventions

The evaluation employed a set of targeting interventions as the lens through which to study the impact of the input subsidies. The conceptual framework of this study builds on existing literature on targeting accuracy and efficiency in development programmes. The term *targeting accuracy* refers to how well the beneficiary group adheres to the targeting criteria. The term *targeting efficiency* refers to how well the beneficiary group was able to make use of the intervention, or package of inputs, in order to increase aggregate food production. In other words, targeting would be considered efficient if the vouchers went to the farmers who are most likely to use the inputs correctly and produce a higher yield than they were previously producing.

Programme targeting is challenging from the perspective of both design (targeting methodology and criteria) and implementation (identification and treatment of target group). Evidence shows that targeted programmes, particularly when pro-poor, often do not reach the intended beneficiaries. A study on the Tanzania's Social Action Fund (TASAF), a community-driven development programme targeting the poor and vulnerable groups, found that TASAF benefits were only mildly pro-poor (Baird, McIntosh and Özler 2011). Similarly, a study on agricultural input subsidy coupons in Malawi by Chirwa, Matita and Dorward (2002) found that vulnerable households, such as those that were poor or with an elderly head, were less likely to receive the fertiliser subsidy. These studies indicate that even when targeting criteria are defined, they are often operationalised in a way that does not result in the intended beneficiary group. It is, therefore, important to verify the accuracy of targeting in order to evaluate the effectiveness of an intervention.

Another body of literature explores the topic of targeting efficiency to understand how to target those who will make best use of the programme. This literature supports the idea that targeting based on observables is imperfect, and that community or self-selection may allocate resources more efficiently because the selectors have access to private information that is a strong predictor of intended programme outcomes. A study of the 1990s' Albanian social safety net programme, Ndihme Ekonomika, found that the programme, which relied on local authorities to select low-income households in need of the transfer, was well-targeted towards the poor (Alderman 2002). The results implied that community administrators had knowledge

about potential beneficiaries that would not necessarily manifest in surveys typically used for targeting. A recent paper on the allocation of land use subsidies in Malawi found that resources were better allocated to the recipients who maximised programme benefits when targeting was implemented through self-selection rather than a lottery (Jack 2013). These studies highlight the tension between designing the targeting methodology in such a way that criteria are well-defined and observable, and, therefore, verifiable and replicable, and relying on local actors to target based on unobservables that may be most relevant to the intended outcomes of the project.

Both schools of thought around targeting are relevant to NAIVS, as the programme has accuracy and efficiency goals. It is intended to produce a macro-level improvement in the supply of food grains by directing the input subsidy to farming households that are not currently taking advantage of improved seeds and chemical fertiliser. While the programme is not intended to be pro-poor, it is designed to focus on smallholder farmers cultivating no more than one hectare of land, and to give priority to women farmers and those who are not already in the habit of using improved inputs. There is an inherent tension between providing inputs to the farmers who could put them to the best use and achieve the largest increases in yields (i.e., better-performing farmers who are likely to be wealthier, cultivating larger plots of land, and/or already familiar with applying improved inputs), and reaching the intended beneficiary group, which may not see high yield growth due to lack of experience with the inputs or simply due to the fact that they may not be the best farmers to begin with. This evaluation's targeting study was intended to look at both questions:
(1) targeting accuracy: when left to the VVC, does NAIVS reach the intended beneficiary group; and (2) targeting efficiency: does the VVC target the vouchers more efficiently to achieve intended outcomes, relative to a lottery?
The underlying hypothesis is that the VVC has private information about farmers' ability to pay for the top-up and skill in using the inputs, which leads to more efficient allocation through normal VVC procedures.

2.2.1 Interventions

In order to generate a valid counterfactual for the treatment group, and to ensure that the preferential targeting of women farmers was successfully implemented, the programme rollout also included an intervention aimed specifically at assessing the targeting accuracy in three regions (Morogoro, Arusha and Kilimanjaro). The interventions were possible only in these three regions given their later planting season. A 2x2 factorial design was used to assign villages to a combination of two targeting interventions:
(1) public meetings for eligibility identification; or
(2) public lottery for beneficiary selection.
The interventions are described below:

Intervention #1: eligibility identification: The goal of this intervention was to enforce the NAIVS eligibility criteria for the selection of new voucher beneficiaries and ensure that certain households that are eligible for the programme are not discriminated against. In one set of villages, a two- to three-person team lead by an MAFC representative facilitated a village assembly meeting to identify households that were eligible to receive the new 2010 NAIVS voucher allocation according to the programme criteria. The meeting was conducted by the VVC with the support of the MAFC agent, who was knowledgeable about the NAIVS eligibility criteria; it was open to all village residents. During the meeting, the VVC was tasked with identifying all remaining eligible households who had not yet participated in NAIVS in the

presence of the village assembly. The criteria used to assess eligibility were:
(1) the cultivation of more than one hectare of maize or rice;
(2) priority given to female-headed households; and
(3) the ability to pay the top-up.
Other eligibility criteria outlined in the NAIVS design, such as willingness to follow extension advice and to serve as a role model in the village, would be difficult to gauge and were, therefore, excluded. The stipulated prioritisation of farming households that have not used chemical fertiliser or improved seeds in the past five years was also excluded because it was not communicated to VVCs in the MAFC guidance in 2010. The final eligibility list was be ratified by the village assembly.

Intervention #2: beneficiary selection: The purpose of this intervention was to introduce a fair and transparent process for beneficiary selection, where eligible female household heads receive priority and all other eligible households have an equal probability of selection. In one set of villages, the intervention team facilitated a public lottery to select X_i2010 new voucher beneficiaries from the VVC's list of eligible farmers, where X_i2010 was the number of new vouchers received by village that year. The names of all eligible farmers were placed in a basket and beneficiaries were randomly selected. To enforce the prioritisation of female-headed households, the names of any female-headed households were entered three times in the basket. The final beneficiary list was ratified by the village assembly and publicly displayed.

Overall, in one-half of the villages, a public meeting took place to identify all eligible households, and in another half, a lottery was introduced to randomly select beneficiaries from the list of eligible households. Thus, villages were randomly assigned to four bins (see Figure 1). In the first bin (PL villages), eligible households were identified via a public meeting (P) and beneficiaries were selected using a lottery (L). In the second bin (VL villages), the VVC identified eligible households (V) and a lottery (L) was implemented to select beneficiaries. In the third bin (PV villages), households were identified using a public meeting (P), but the VVC decided who the beneficiaries would be among the eligible households (V). Finally, in the fourth bin (VV villages), the normal selection process took place whereby the VVC identified eligible households (V) and selected beneficiaries (V). This design allows us to not only estimate the accuracy of the VVC targeting but also to cleanly measure the impacts of the programme (see Chapter 3 for more detail on evaluation design).

Figure 1: Factorial design of the targeting interventions implemented in Arusha, Morogoro and Kilimanjaro

PL Villages Public meeting + lottery	**VL Villages** Lottery only
PV Villages Public meeting only	**VV Villages** No targeting intervention (control)

2.2.2 Theory of change

The theory of change for the targeting interventions is provided in Appendix B and explains the hypothesis driving the targeting study. This underlying hypothesis is that VVCs are able to target better-performing farmers based on knowledge of unobservables, and so beneficiaries identified by VVCs' standard procedures will produce higher yields than non-beneficiaries in the same villages. In villages where eligibility criteria are strictly enforced and beneficiaries are

selected by lottery, less skilled farmers may be selected as beneficiaries, and so the difference in yields between beneficiaries and non-beneficiaries in these villages will be smaller. Similarly, yields produced by these villages' beneficiaries will be lower compared to those of the beneficiaries in villages where the VVCs selected freely. In PL villages, we expect farmers cultivating large plots of land (over one hectare) and who have previously used improved inputs to be excluded; therefore, beneficiaries are likely to include a mix of skilled and less skilled farmers. We expect to see an insignificant difference in yields between beneficiaries and non-beneficiaries, non-compliance of prescribed input use among beneficiaries and perhaps some evidence of sale or sharing of inputs by beneficiaries. In VL villages, eligibility criteria may not have been adhered to as strictly and so the eligible pool may include some of the village's best-performing farmers. The lottery would result in a mix of skilled and less-skilled beneficiaries, but the beneficiary group will likely produce yields somewhat higher than non-beneficiaries, though perhaps not significantly so. We still expect to see some non-compliance in input use and some level of sales or sharing of inputs by beneficiaries. In PV villages, we expect the VVC to select the best-performing farmers as beneficiaries; therefore, the difference in yields between beneficiaries and non-beneficiaries is expected to be significantly large and we do not expect to find evidence of sharing or non-use of improved inputs by beneficiaries. In VV villages, the same results are expected as of PV villages, but to a more significant degree, as VVCs had free reign to target the best farmers in the village.

2.2.3 Key indicators

The key indicators included in the evaluation to analyse the expected outcomes of the targeting intervention include:

- Eligibility status
- Beneficiary status
- Use of vouchers
- Plot-level input use
- Maize and rice yields
- Sharing or sale of inputs

3. Evaluation design

3.1 Overview

Evaluation activities took place between December 2010 and December 2012. To answer the core evaluation questions related to programme impacts on agricultural production, nutrition and household welfare, including heterogeneous impacts by the gender of the household head, we constructed a two-year panel survey with approximately 2,000 households (referred to as the household survey), half of which were new NAIVS beneficiaries in the 2010–2011 planting season. The other half were identified as eligible for NAIVS in 2010–2011 but had not yet participated in the programme.

3.2 Study population

The overall sample for the impact evaluation was selected from high-potential maize or rice growing regions in the NAIVS programme area. Tanzania's administrative areas are divided into regions, districts, wards, villages and sub-villages or hamlets, each of which is a subset of the previous unit. The sample regions, noted below, spanned the southern highlands, northern

highlands and part of Tanzania's central zone. Three of these regions, Arusha, Kilimanjaro and Morogoro, are distinct in that they experience bi-modal rainfall patterns, which result in a later start date for the main planting season as compared to other areas with uni-modal rainfall. Bi-modal rains meant that planting took place later than regions with uni-modal rains and allowed us time to conduct the targeting interventions to select voucher beneficiaries in these areas. Given that the targeting interventions were the mechanism through which to randomise treatment and establish a valid counterfactual, the impact evaluation was intended to focus on the regions where the interventions took place. The five additional regions were included in the sample to provide additional context to the study from the southern highlands, which comprise the 'grain belt' of Tanzania.

Table 2: Regions/districts

Northern highlands	**Southern highlands**
Arusha (Meru)^	Iringa (Kilolo & Njombe DC)
Kilimanjaro (Same)^	Mbeya (Mbeya rural & Mbozi)
	Rukwa (Sumbawanga DC)
Central zone	Kigoma (Kasulu)
Morogoro (Ulanga)^	

Note: ^Experiences bi-modal rainfall and, therefore, a later planting season.

3.3 Sample selection

The later start date for the planting season in Arusha, Kilimanjaro and Morogoro allowed for the village-level targeting interventions to be conducted in the bi-modal regions *prior to* the distribution of the 2010–2011 vouchers. Each region was assigned 40 villages, with 10 households per village. The remaining 80 villages of the 200 total were spread evenly across the other five uni-modal regions, with 16 villages per region and 10 households per village. Districts in each sample region were assigned weights based on the total number of eligible farming households and a rule was established to randomly sample one in four districts per region. Using this rule, larger districts had a higher probability of selection and, in regions with up to four districts total, one sample district was selected. This resulted in one sample district for all regions, except Iringa and Mbeya, where two sample districts were selected (the 16 villages allocated per region were divided evenly between the two districts). In the bi-modal regions, where the allocated sample was much higher, villages were randomly selected across all wards. In the other five regions, wards were randomly selected (the number of wards sampled was equal to the number of villages that would be sampled) in each sample district and one village was randomly selected within each sample ward. The village sampling frame consisted of villages that would be participating in NAIVS in the 2010–2011 planting season and would be receiving an increase in the number of vouchers distributed relative to last year. The second condition was put in place to ensure that there would be scope for new beneficiaries to participate (and sample from) in addition to the previous beneficiaries who would be receiving vouchers for the second or third time. Villages were randomly selected from this sampling frame. During implementation of the targeting interventions, it was discovered that the VVC had already selected 2010–2011 beneficiaries prior to the arrival of the intervention teams in the majority of villages in Ulanga (Morogoro) and Same (Kilimanjaro). It was not possible to redo beneficiary selection in those villages; therefore, the targeting interventions could not be implemented there. In order to maximise the sample in areas where the interventions were implemented successfully, six of the 40 villages assigned to Kilimanjaro were reallocated to Arusha, allowing for six additional villages in Meru district to be randomly selected for the baseline. This brought Arusha's total to 46 and Same's to 34.

The follow-up study revisited the same households and villages as the baseline; however, in the follow-up survey, given difficulties in implementing the targeting interventions in two of the intervention regions, the sample size per region was altered in order to enhance the sample size where targeting interventions had been successful (see 3.7, 'Changes to the sample', below for additional details). Ultimately, the analysis focused on a subset of the overall sample.

3.4 Sample size and power calculations

Power calculations were conducted only for the household survey, which will be used to determine programme impacts. Based on calculations using both the 2003 Agricultural Census and 2008 National Panel Survey data, the necessary sample size for the household study was estimated to be between 1,040 and 1,400 households, or approximately 100 to 140 villages, with 10 observations (farming households) per village. This sample size allows the detection of a change in yields of at least 34 per cent across treatment and control groups, with a 5 per cent probability of Type I errors and 10 per cent probability of Type II errors. In order to maximise power in comparisons across treatment and controls, approximately 120 villages were originally sampled in the three regions where targeting interventions were conducted and 60 villages were sampled in the other regions. This amounted to a sample size of 2,000 households in total, with 1,200 in the targeting intervention regions.

Additional power calculations were conducted following the baseline survey once we learned that the targeting interventions had not been successfully implemented in two of the three northern highland regions (see 3.7, 'Changes to the sample'). These revised calculations revealed that if we doubled the sample size in Arusha (from 460 households to 920 households), we would have the power to detect a 30 per cent increase in yields, assuming a conservative intra-class correlation of 0.1 and 46 clusters. Given that other data collected on the voucher programme predicted yield changes of more than 200 per cent, we were comfortable with this change and indeed expanded the sample in Arusha.

3.5 Data collection

In preparation for the evaluation, we reviewed existing data sets, such as the Tanzanian Agricultural Census and Living Standards Measurement Survey. However, we found that they did not provide the geographic scope, plot-level detail and programmatic questions required to comprehensively evaluate a programme of this nature. We, therefore, created a set of four questionnaires (described below) that were administered across the 200 villages at baseline, three of which were repeated at follow-up. Baseline surveys were administered in January and February of 2011; follow-up data collection took place in July and August of 2012. All data collection activities were conducted by an in-country survey firm, Research on Poverty Alleviation (REPOA), in partnership with researchers from MAFC, the World Bank and the University of California, Berkeley.

3.6 Data collection instruments

Table 3: Summary of quantitative data collection tools

Survey	Brief description on instrument and population	Sample size	
		2010–2011	2011–2012
Household survey	Detailed questionnaire with modules covering household-level demographics, food security, plot-level crop production and marketing, NAIVS implementation and 5 gender-related modules for women dealing with violence, power, relationships and self-reported measures of health. Sample purposefully selected to include approximately one half new beneficiaries and one half eligible non-beneficiaries.	2,000	2,040
Village survey	Short questionnaire to collect village-level data and output indicators on NAIVS implementation. In all participating villages, one village leader sampled.	200	174
Agro-dealer survey	Short questionnaire to collect village-level data on input prices, availability and NAIVS implementation. In all participating villages, 1–2 agro-dealers sampled.	88	45
VVC survey	Short questionnaire on village-level beneficiary selection process. Conducted with 1 male and 1 female VVC member in participating villages.	200	N/A

Together, these data will allow for a comprehensive analysis of NAIVS's impact with gender-disaggregated impacts across all indicators, including plot-level crop effects, household economic impacts, health outcomes, food security and spillovers to the agricultural input and output markets.

Household survey: In each study village, 10 households participated in the household survey, five of which were new beneficiaries and five eligible non-beneficiaries. Eligibility and beneficiary status were defined by the VVCs (with or without the help of the targeting intervention facilitators) and households were sampled using VVC records. In order to investigate gender-related impacts of NAIVS, women were oversampled such that 50 per cent of the household survey sample in each village consisted of female-headed households.

Other surveys: Additionally, a VVC survey looking at voucher operations in the village was conducted with one male and one female member of the VVC, a village survey was conducted with the village executive officer (VEO) to record community-level characteristics, and finally an agro-dealer survey was administered to all agro-dealers working in the village to understand the input supply side of the programme.

3.7 Changes to the sample

In the bi-modal regions, because targeting interventions were conducted as planned only in Arusha, households in the 46 Arusha villages were oversampled (doubled) in the follow-up survey, resulting in a sample of 920 households. Households for the expanded sample were selected using the same mechanism utilised in the baseline; specifically, eligible beneficiaries were defined as those that began participating in NAIVS prior to the 2010–2011 planting season and non-beneficiaries were defined as those that were eligible to participate in 2010–2011 but were not selected. Villages from Morogoro and Kilimanjaro were reduced to compensate for the oversampling in Arusha. This did not significantly change the overall sample size for the northern regions or for the overall household survey. However, increasing the Arusha sample size allowed the analysis to focus on this particular region where the targeting interventions were implemented effectively.

3.8 Framework and methods for analysis

The evaluation was able to exploit an exogenously driven delay in the rollout of the programme to generate comparable treatment and counterfactual groups for estimation of programme impact. Prior to the original implementation of the programme, it was estimated that 2.5 million households were eligible; however, the government distributed vouchers to only 1.5 million households in 2009–2010 and 2 million households in 2010–2011. Thus, each year there have been fewer vouchers distributed than the number of eligible farmers, and we used this shortage to create an appropriate counterfactual, or comparison, group of farmers. Therefore, when estimating programme effects using the core household survey, we are comparing eligible beneficiary households and eligible non-beneficiary households within the same village. The treatment group is comprised of farming households that began receiving the subsidy for the first time in the 2010–2011 planting season and the comparison group consists of eligible farming households in the same village that had never received the subsidy.

However, because in most of these villages the assignment of vouchers was done by the VVC, it is unclear whether beneficiaries and non-beneficiaries are comparable. We, therefore, focus the impact analysis on the subset of villages that were randomly assigned the targeting interventions to generate the appropriate treatment and counterfactual groups, and to explore the effect of programme targeting by the VVC. By virtue of the lottery, the identification of beneficiaries in villages where the lottery was implemented is random. Indeed, we can test the presence of selection on observable and unobservable characteristics by comparing outcomes in villages with the lottery to villages without. When the outcome variable is continuous, we used the following ordinary least squares (OLS) regression:

$$Y_{ij} = \alpha + \beta_{Ben} Benef_{ij} + \beta_{PL} PL_j + \beta_{VL} VL_j + \beta_{PV} PV_j + \beta_{BenxPL} Benef_{ij} x PL_j + \beta_{BenxVL} Benef_{ij} x VL_j + \beta_{BenxPV} Benef_{ij} x PV_j + \varepsilon_{ij,}$$

where Y_{ij} is a given outcome for individual i in targeting village j (PL, VL or PV, see Figure 1), $Benef_{ij}$ is a dummy variable that takes value 1 if individual i in village j was selected as a beneficiary, dummies, PL_j, VL_j and PV_j indicate whether village j is under treatment PL, VL or PV. The term ε_{ij} is a mean-zero error and because the unit of randomisation is the individual, standard errors are clustered at this level when multiple observations per individual are reported.

When the dependent variable was binary, we used an OLS specification to ease the interpretation of the coefficients,[2] and when the dependent variable was truncated from the left, as in the kilogram of fertiliser used in the field, Tobit specifications of the equation above were used. When we assessed nutrition outcomes, we used a linear probability model and logistic regression to examine the association between beneficiary status and food insecurity, and OLS regression to explore the association with dietary diversity.

The coefficients of interest are the β coefficients and measure the impact of the programme. In particular, β_{Ben} measures the impact of the programme on beneficiaries in VV villages relative to non-beneficiaries in VV villages. Coefficient β_j measures the impact of the programme in non-beneficiaries in villages assigned to treatment *j,* compared to non-beneficiaries in VV villages. Finally, the coefficient β_{Benxj} measures the differential impact of the programme on beneficiaries in village of treatment *j* relative to beneficiaries in villages assigned to VV treatment. In order to calculate other comparisons of interest, contrast statements (e.g., lincom) were used to estimate β coefficients and assess statistical significance.

To measure the impacts on women compared to those of men, we run separate regressions among male-headed and female-headed households. For select outcomes (such as household decision-making autonomy), we restrict the regressions to the sample of women only.

Overall, this analysis framework permits comparison of two similar groups that differ only with respect to their exposure to the programme. As a result, differences in outcomes can be causally attributed to the programme and not to other underlying or unmeasured causes. Impact results from this analysis are provided in Chapter 5. However, results of the NAIVS implementation assessment follow in the next chapter.

4. Programme implementation

The household, VVC and agro-dealer surveys collected detailed information on various NAIVS procedures from the perspective of farmers, VVC members and agro-dealers. This data, taken in conjunction with anecdotal evidence from field visits, reveals that programme implementation tends to vary somewhat by district and in some cases by village. This section looks at issues of programme awareness and procedures, VVC operations, agro-dealer operations and targeting.

4.1 National Agricultural Voucher Scheme awareness and procedures

Household and VVC baseline survey respondents were asked detailed questions about their awareness relating to the rules and procedures of NAIVS. On the whole, the data revealed some gaps in information both among households and VVC members. General awareness of the voucher programme was high, as 76 per cent of households had heard of it at baseline (Table 4). However, there was some indication that households were not universally clear about the procedures for programme eligibility. The most notable knowledge gap related to the three-year rule for receiving the subsidy; only 6 per cent of household survey respondents knew of the limit, while 25 per cent thought the subsidy would be provided indefinitely. In the

[2] We reran all specifications with binary dependent variables using a Probit model and the results were very similar. None of the significance levels change. Therefore, we present primarily the OLS models to ease interpretation of results.

follow-up survey, general knowledge of the programme had improved; however, there was still limited knowledge of certain programme aspects, for example, the three-year duration of the programme, among households.

Table 4: Awareness about NAIVS in the baseline VVC, and baseline and follow-up household surveys

	Households				VVC members	
	Baseline		Follow-up		Baseline survey	
	N	%	N	%	N	%
Awareness of existence of NAIVS	2,000	76	2,040	94		
Awareness of 3-year duration of subsidy	1,520	6	2,040	28		
Awareness of own eligibility status	1,520	43	2,040	68		
Awareness of eligibility criteria	1,520	46	--	--	420	85
Knows the responsibilities of the VVC					420	86
Knows of beneficiary selection responsibility					361	46
Knows of voucher distribution responsibility					362	72

Note: Unlike the baseline survey, there was no skip pattern in the follow-up survey that prevented individuals who had not heard of NAIVS to try to answer subsequent questions related to programme details.

Household survey respondents were asked a module of questions relating to various aspects of programme implementation in order to understand how NAIVS operated in different villages. Across implementation indicators, a selection of which is reported in Table 5, adherence to guidelines improved with each year of the programme. Still, certain procedures, such as assessment of the top-up criterion and identification of the eligible, were not followed uniformly. This could have been due to a lack of awareness on the part of the VVCs and village leaders or villages may have developed their own voucher management systems to best suit their environment.

Table 5: Household baseline survey—NAIVS implementation

	2008	2009	2010
Asked if could afford top-up	7.8%	16.5%	24.5%
Stated a village meeting took place to identify the eligible	51.1%	55.5%	57.4%
Knew of NAIVS complaint procedure existing in village	12.0%	14.5%	15.4%
For respondents in southern regions, where previous beneficiaries were included in the sample:			
Stated they paid something to become voucher beneficiaries	2.4%	2.9%	2.7%
Had physical possession* of vouchers	86.1%	85.2%	89.7%

Note: *The alternative to possession was often a system where VVCs would store vouchers and oversee redemption so that farmers handled and signed vouchers at redemption only.

4.2 VVC operations

The VVCs were the designated focal points for voucher activities in NAIVS villages. They were designed to consist of six village members (three men and three women) elected by the village assembly to oversee the voucher programme at the village level. As reported in Table 6, 84 per cent of the 200 VVCs surveyed followed guidelines and had three male and three female members, and most were nominated by the village government or a village assembly. While the original intent was to have a standing VVC that would continue from year-to-year, in reality, 34 per cent of the 200 VVCs interviewed had changed membership since the beginning. This could pose problems for consistent application of NAIVS rules and procedures, particularly in terms of recordkeeping to prevent duplicate beneficiary households. During site visits, a common complaint amongst VVC members was a lack of compensation for their time and efforts. In the survey, 9 per cent of respondents stated that the VVC received remuneration between TZS1,000 and 15,000 (approximately USD 0.60 to 10). If the VVC is to be relied upon for organised beneficiary selection as well as monitoring of voucher redemption and input use, a small amount of compensation may incentivise members to fulfil each of these roles.

VVC members were asked to describe the process by which eligible households were identified and beneficiaries were selected. About three fourths of respondents stated that they received recommendations of eligible farmers from the village, hamlet or 10-cell leaders both in 2009–2010 and 2010–2011 (Table 5). The majority of respondents (over 80 per cent) notified eligible farmers and beneficiaries through a village assembly in both years. Only 20 per cent of VVCs surveyed made the beneficiary selection decision, while 40 to 47 per cent let village leaders decide, and 33 to 40 per cent left the decision to the village assembly. Having village leaders select beneficiaries may result in a loss of accountability that exists in community-based targeting; however, based on targeting results this does not seem to have been the case.

Table 6: VVC structure

Membership has changed since formation	34%
Number of times changed:	
1–3	45%
4–6	46%
VVC had 3 male + 3 female members at formation	84%
Process of VVC member selection:	
Elected	21%
Nominated by village assembly	30%
Nominated by village government	48%
VVC members receive compensation	9%

Table 7: VVC general procedures

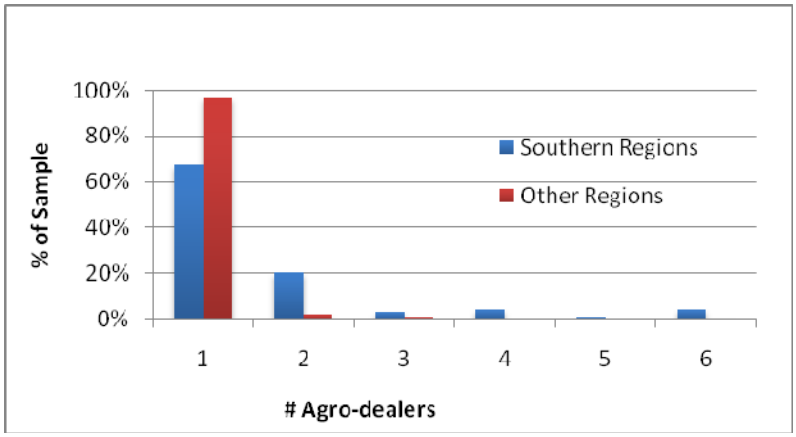
	2009–2010	2010–2011
Most common method to identify eligible farmers:		
VVC gets recommendation of the eligible from village, hamlet or 10-cell leaders	76%	75%
Most common method to notify eligible farmers:		
Village assembly announcement	87%	82%
Method of selecting beneficiaries from among the eligible:		
Village leaders decided	47%	40%
VVC decided	20%	20%
Village assembly decided	33%	41%
Method of notifying beneficiaries:		
Village assembly announcement	86%	84%
List publicly displayed	9%	14%

4.3 Agro-dealer operations

A key component to the operation of NAIVS is the expansion of the agro-dealer network to ensure adequate stocks and timely delivery of inputs in villages. Despite the incentives that the NAIVS provides to encourage agro-dealers to start or spread operations across villages, 67 per cent of VVCs surveyed in the south and 97 per cent of villages in other regions reported having just one agro-dealer serving the village at the time of survey. In some cases, districts assign one agro-dealer to a village to ensure that all villages have a dealer selling to them. In other cases, methods for encouraging competition amongst agro-dealers should be considered, such as another round of CNFA training to add to the pool of certifiable agro-dealers. Issues such as credit from wholesale input suppliers and manufacturers or the timeliness of voucher cash flows may also affect agro-dealers' decisions to expand operations to voucher villages. On average, during the 2009–2010 planting season, agro-dealers had to wait 14 days to receive money from the bank after submitting redeemed vouchers. Late or slow payment was a common complaint amongst agro-dealers and runs the risk of deterring future participation in the programme.

While current agro-dealer operations seem to have met the demand for voucher input purchases, the sustainability of these operations is not guaranteed. Seventy-one per cent of agro-dealers surveyed became involved in their agro-dealer business in 2008 or later, which implies that they likely started in response to the NAIVS. In addition, 70 per cent of household survey respondents stated that only farmers with vouchers could buy inputs from the voucher-certified agro-dealer in their village in 2010–2011, which means that these agro-dealers are counting solely on subsidises purchases for their operations. There is, therefore, a definite risk of these new agro-dealer businesses closing or downsizing operations once the subsidy ends, so it is important to consider how to encourage greater competition and long-term profitability for dealers to operate in absence of the vouchers.

Chart 1:Agro-dealers selling in village, as reported by VVC



In terms of voucher sales practices, as reported in Table 5, it seems most agro-dealers do not stray from the types and quantities of inputs as prescribed by the vouchers when faced with buyers who want to use the vouchers for different purposes. When posed with hypothetical scenarios about beneficiary redemption, over 70 per cent of agro-dealers stated that farmers could only buy the specified type and amount of the input listed on the voucher. In cases where the buyer cannot afford the top-up payment, however, there is some indication that dealers may be flexible, as only 49 per cent said the buyer would not be able to purchase the inputs without the full top-up. These results provide a sense of agro-dealer behaviour, but do not give an indication of the number of vouchers that are being redeemed incorrectly.

Table 8: Voucher-related operations of agro-dealers

What happens if farmer does not want voucher-specified input?	
Can only purchase specified input	74%
Can purchase other inputs instead	18%
What happens if the farmer only wants to buy part of the voucher-specified quantity?	
Can only purchase specified amount	85%
Can purchase less than the specified amount at a discount	2%
What happens if the farmer does not have the full top-up amount?	
Cannot purchase inputs using the voucher	49%
Can purchase a lesser quantity using the face value of the voucher	13%

4.4 Targeting

The household survey collected information related to the NAIVS eligibility criteria in order to assess the adherence to programme guidelines in beneficiary targeting. Given the high costs associated with providing subsidies, a universal NAIVS programme for all farmers would not be feasible and so beneficiary targeting criteria were established to identify farming households that would make best use of subsidised inputs. Eligibility for the voucher programme was to be based on two hard criteria:

(1) the household was cultivating no more than one hectare (approximately 2.5 acres) of maize or paddy; and

(2) the household could afford the top-up payment for the input package (approximately TZS55,000, but variable by region and year).
In cases where the number of eligible households was larger than the number of vouchers, priority was to be given to female-headed households and farmers who had not been using improved inputs in the five years prior to the start of the NAIVS.

Table 9 reports listing survey baseline results relating to targeting from the full household sample and that of two geographic zones (northern and southern regions). Based on respondents' self-reported areas of maize cultivated in 2009–2010 and ability to pay the top-up during the 2010–2011 planting season, approximately 52 per cent of the overall sample and 62 per cent in the Arusha sample were eligible for maize vouchers. Overall, 56 per cent of the full sample and 65 per cent of the Arusha sample were technically eligible to receive either maize or paddy vouchers. Given that the entire survey sample included households that had been identified as eligible by the VVC, these results indicate that approximately 44 per cent of households in the sample had been incorrectly identified as eligible. The land criterion was a larger limiting factor for eligibility than the top-up requirement, as 69 per cent met the land requirement for maize, while 78 per cent of respondents stated they could afford the top-up payment. The following sections discuss each of the targeting criteria and priorities in more detail.

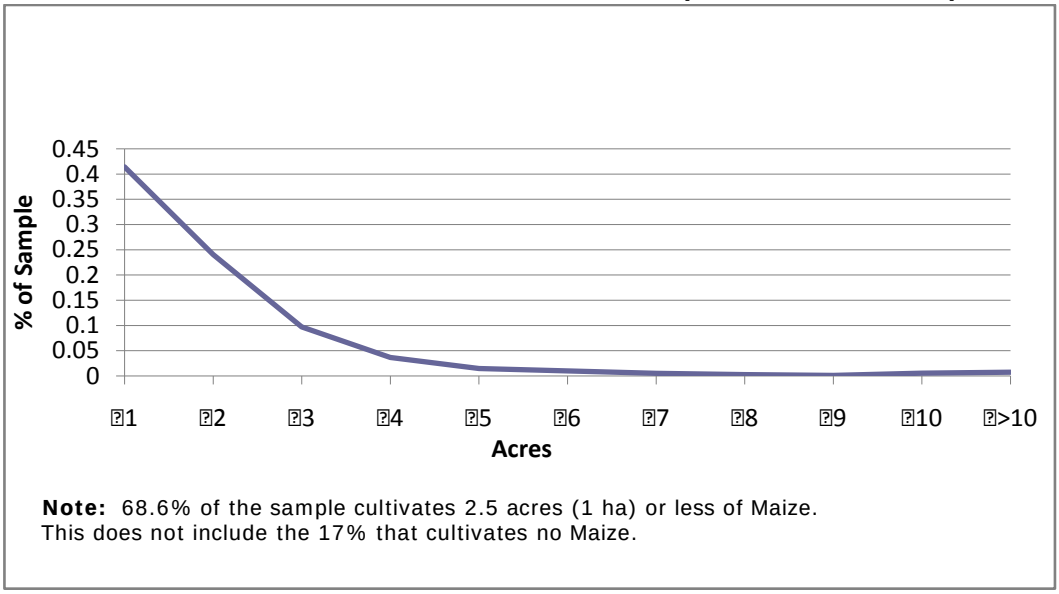
Table 9: NAIVS eligibility (for 2010–2011 voucher cycle)

	N (all regions)	All regions	Arusha only
Could afford the top-up in 2010–2011	1,343	77.7%	84.5%
Cultivated less than 1 ha of maize in 2009–2010	2,000	68.6%	77.6%
Cultivated less than 1 ha of paddy in 2009–2010	2,000	16.6%	3.0%
Eligible for maize voucher (land + top-up criteria)	1,343	52.1%	62.2%
Eligible for paddy voucher (land + top-up criteria)	1,343	10.7%	5.0%
Eligible for any voucher (land + top-up criteria)	1,343	56.1%	64.7%

4.4.1 Land

Taking a closer look at the land criterion for voucher programme eligibility, it is evident that it has been difficult to enforce. To establish land eligibility for the 2010–2011 voucher cycle, the survey recorded households' areas of crop cultivation from the previous planting season. Households cultivating neither maize nor paddy were not eligible for that particular type of voucher. While 31 per cent of the survey sample cultivated more than one hectare of maize in 2009–2010 (and were therefore ineligible), the majority of this ineligible group cultivated less than two hectares of maize (approximately five acres), which implies that they were still farming relatively small plots of land (see Chart 2).

Chart 2:Land area under maize cultivation (2009–2010 season)



Note: 68.6% of the sample cultivates 2.5 acres (1 ha) or less of Maize. This does not include the 17% that cultivates no Maize.

4.4.2 Top-up

In order to determine a household's ability to pay the top-up, the survey asked respondents to self-report whether they could afford the expected top-up amount and also recorded information related to household economics to evaluate whether the household was credit constrained.

Among the baseline household survey sample of eligible households, 78 per cent stated they were able to afford the top-up in 2010–2011. However, from the perspective of savings and credit, 78 per cent of the sample revealed themselves to be credit constrained. This would imply that a smaller percentage of the sample than self-reported is actually able to afford the top-up. However, 94 per cent of respondents who received vouchers in 2010–2011 paid for the voucher inputs using cash (Table 10), suggesting that, somehow, farmers are able to find ways to pay for the top-up.

Table 10: Source of payment for the top-up among baseline household survey respondents

	2009–2010	2010–2011
N	269	347
Cash	79%	94%
Another voucher	2%	0%
Loan	1%	1%
Other	18%	5%

4.4.3 Networks

With community-based targeting comes the risk of elite capture, whereby village members with positions of power can exert influence over the beneficiary selection process and, in the case of the voucher programme, direct vouchers to friends or relatives in their own networks. The household survey investigated whether connectedness to positions of power within the voucher programme had any impact on becoming a beneficiary by asking respondents to

characterise their relationship, if any, to the hamlet chairperson, village executive officer and members of the VVC. Respondents who held any of the positions or were related by blood to any position holders were considered 'related'. Those who knew the position holders and had regular social contact with them (defined as having meals in each other's homes) were considered 'connected'. Regressing the dummy variable for beneficiary status on various demographic indicators did not reveal a clear bias towards connected farmers for beneficiaries (Table 11). It is evident that being literate, owning larger areas of land or having greater wealth were correlated with eligible households becoming beneficiaries, but neither being related nor connected to positions of power resulted in statistically significant coefficients. Being a male household head results in a slightly lower likelihood of being a beneficiary, which supports the idea that prioritisation of female-headed households was being upheld. In terms of networks, however, the survey does not support the idea that vouchers went to the most connected farmers.

Table 11: Regression—determinants of beneficiary status among eligible households

Variable	Coefficient	Standard Error	
Male household head	-0.04	0.02	*
Literate household head	0.11	0.03	***
PCA –– housing wealth	0.03	0.01	***
Cultivable land area owned	0.003	0.00	***
Related to district, ward, village or hamlet officer	0.09	0.06	
Connected to district, ward, village or hamlet officer	-0.08	0.07	
Constant	0.44	0.03	***
Note: N = 1,934; R^2 = 0.02. P-values = *** < 1%, ** < 5%, * < 10%.			

4.5 National Agricultural Voucher Scheme awareness and procedures from follow-up

Table 12 reports the responses to a module in the follow-up household questionnaire. Column 1, similar to the results from the baseline, suggests that most households, around 80 per cent, are aware of the programme. Of course, virtually all beneficiaries are aware of the programme.

Column 2 reports the likelihood that respondents attended a meeting to identify eligible farmers. As can be seen from the p-values below, there are no differences in the probability of attending the meeting between those that would ultimately become beneficiaries and those that would not. In contrast, in VL villages, only beneficiaries are more likely to recall a meeting. Since public meetings were supposed to happen in PL and PV villages, this pattern is consistent with meetings taking place in PL and PV villages but not in VL. We note that there is no difference in the prevalence of attending a meeting between beneficiaries and non-beneficiaries in VV villages either. Column 3 asks about whether the respondent was asked if he or she could afford the top-up. This is important because the VVC would report to us anecdotally that they 'knew' who could and could not pay, and only offered the vouchers to those that could pay. Put differently, the ability to pay could be used as an excuse to target the vouchers to specific individuals. In PV villages (but not in VV), we find that only beneficiaries tended to be asked. In PL villages (but not PV), both beneficiaries and not were more likely to be asked, as intended.

Columns 4 and 6 report whether the beneficiary acknowledged receiving the vouchers. Recall that the definition of beneficiary we use here is taken from the administrative data, so there could be differences between the administrative data and the actual voucher allocation. These columns suggest that the mismatch is non-trivial. Indeed, around 20 per cent in 2010–2011 and 24 per cent in 2011–2012 of individuals classified as non-beneficiaries in 2010–11 received vouchers. More problematic, only around 45 per cent of beneficiaries reported having received the vouchers. Interestingly, the village type does not predict where the mismatch is likely to occur, but the magnitudes are somewhat large.

Finally, column 5 reports the satisfaction of villagers with the selection of beneficiaries. We find that in all villages except for PL villages, beneficiaries are, not surprising, more satisfied with the selection than non-beneficiaries. However, in PL villages both beneficiaries and non-beneficiaries are equally satisfied, suggesting that the lottery could be seen as fair by everyone. We note that, in VL villages, a lottery was also conducted and yet it seems that this pattern is not found in the data.

Table 12: NAIVS awareness and implementation among Arusha households in the follow-up survey

	Awareness of NAIVS programme (1= yes)	**Attend meeting where the eligible were identified in 2010–2011**	**Asked about ability to pay in 2010–2011**	**Household received vouchers in 2010–2011**	**Satisfaction with selection of beneficiaries in 2010–2011**	**Household received vouchers in 2011–2012**
	(1)	(2)	(3)	(4)	(5)	(6)
Beneficiary	0.092**	-0.188	0.065	0.253***	0.197***	0.171***
	(0.042)	(0.195)	(0.052)	(0.060)	(0.064)	(0.062)
PL village	0.063	-0.232	0.044	0.113*	0.217***	0.092
	(0.044)	(0.194)	(0.052)	(0.058)	(0.064)	(0.061)
VL village	0.033	-0.303	0.027	0.083	0.168***	-0.007
	(0.046)	(0.221)	(0.050)	(0.057)	(0.065)	(0.057)
PV village	0.023	-0.359*	-0.013	0.032	0.161**	-0.007
	(0.048)	(0.209)	(0.049)	(0.056)	(0.066)	(0.058)
Beneficiary x PL	-0.033	0.396	-0.035	-0.033	-0.228***	-0.048
	(0.052)	(0.241)	(0.074)	(0.087)	(0.086)	(0.088)
Beneficiary x VL	0.004	0.575**	0.051	0.091	-0.085	0.086
	(0.053)	(0.277)	(0.076)	(0.085)	(0.085)	(0.086)
Beneficiary x PV	0.011	0.335	0.028	0.096	-0.079	0.009
	(0.055)	(0.282)	(0.074)	(0.086)	(0.086)	(0.088)
Constant	0.848***	0.667***	0.152***	0.190***	0.533***	0.238***
	(0.035)	(0.162)	(0.035)	(0.038)	(0.049)	(0.042)
Obs.	920	130	920	920	920	920
R-squared	0.0312	0.0726	0.0132	0.101	0.032	0.043
Mean of dep. var. in VV	0.895	0.531	0.186	0.323	0.636	0.327
P-val. of beneficiary = Non-beneficiary in…						
PL village	0.063	0.145	0.569	0.000	0.586	0.051
VL village	0.004	0.053	0.035	0.000	0.043	0.000
PV village	0.004	0.472	0.074	0.000	0.042	0.004

5. Impact results

5.1 Impact evaluation sample

Table 13 summarises basic demographic characteristics of the Arusha sample from the baseline (2011) household survey. Since the evaluation had to be limited to Arusha (Meru district), it is important to consider the characteristics of Arusha in the context of the external validity of the study. As has already been noted, Arusha is one of the few regions in Tanzania that experiences bi-modal rainfall patterns and where farmers may choose to plant in both the short and long rains (the surveys focused on the long rains). While the Arusha region comprises parts of the northern highlands as well as arid lands, Meru district is in the highlands and home to both maize and rice cultivation. Anecdotal evidence from field visits and MAFC indicate that Meru district is one of the more organised and well-managed of the districts participating in the NAIVS. This may explain why the targeting interventions were only implemented successfully here.

Table 13: Household survey sample (Arusha) at baseline

Beneficiaries	**52%**					
	N	**Mean**	**SD**	**p10**	**p50**	**p90**
Age of HH head	440	49.1	15.9	30.0	47.0	73.0
Sex of HH head: male	460	0.6	0.5			
Education of HH head: none	441	19%	0.4			
Literacy of HH head: read and write	441	75%	0.4			
Total land area owned (acres)	460	2.6	2.8	1.0	2.0	4.5
Cultivable land area owned (acres)	460	2.1	2.0	0.5	1.5	4.0
Years respondent has been farming	460	24.1	16.2	5.0	20.0	48.0
Monthly HH expenditure (TZS)	456	156,080	147,506	45,000	113,125	313,100
Annual HH income (Dec. 2010) (TZS)	131	1,095,984	1,481,404	50,000	600,000	2,400,000
Annual HH savings (Dec. 2010) (TZS)	455	63,880	311,637	0	0	100,000

Note: The household survey sample was about evenly split between male- and female-headed households by design, with 60 per cent of households headed by a male.

5.2 Impact analyses

As described above, the targeting interventions were successful in Arusha and the sample was subsequently doubled in this region for the follow-up (2012) household survey. Therefore, we have focused our impact analyses to households in this region and all results presented below, with the exception of the results on targeting, are from this sub-sample.[3] The analysis presents intention to treat estimates; while no individuals refused the vouchers, some did not redeem all three vouchers and, therefore, can be considered to have not received full treatment. Table 14 compares select descriptive and socioeconomic characteristics of households in Arusha to those in the rest of the country using the household survey. The table highlights general differences between households in Arusha as compared to households in the remaining geographic regions. For example, across nearly all indicators, including personal demographics, occupation and assets and landholdings, there were statistically significant differences between Arusha and non-Arusha response distributions. Households in Arusha had less land than households in the rest of the country, were less likely to work on their own farm and were more often illiterate, although their houses were more often made of higher quality materials. These results imply that the findings from the impact evaluation may not be generalisable to the rest of the country.

[3]Even though the targeting interventions were successfully implemented in only one region, our collaborators at the World Bank and MAFC felt that it was important to continue to assess programme impacts countrywide in order to be able to assess regional variation programme targeting, implementation and impacts. We recognise that specific analytic methods will need to be employed to overcome the selection bias inherent in these regions, particularly the south, in order to analyse findings in these regions. As a result, we focus our impact analyses to the region (Arusha) where the targeting interventions resulted in random allocation of the voucher and we, therefore, have the potentially cleanest estimates of programme impact.

Table 14: Select descriptive and socioeconomic characteristics of follow-up (2012) household survey respondents, stratified by location

	Location		Sig. using χ^2 or t-test
Household characteristics	**Arusha**	**Rest of sample**	
Male-headed household	54%	53%	
Age of household head (median)	48.7	45.8	***
Household size (median)	5.0	5.0	
Marital status of household head:			**
Married (monogamous)	54	57	
Married (polygamous)	8	4	
Informal union	1	4	
Divorced	1	1	
Separated	7	8	
Widowed	23	20	
Never married	5	6	
Primary occupation:			***
Paid employee	4	2	
Self-employed	1	0	
Unpaid family helper (ag)	74	65	
Works on own farm	17	28	
Literacy:			***
Read only	2	2	
Write only	1	1	
Both	74	83	
Neither	22	13	
Housing construction:			***
Walls (burnt brick or cement)	12	8	
Roof (corrugated tin)	83	61	
Floor (cement)	39	28	
Assets:			**
Radio	68	64	
Mobile phone	61	47	
Refrigerator	2	1	
Electric/gas stove	2	1	
Bicycle	2	1	
Motorbike	7	6	
Motor vehicle	32	29	
Total area land owned	4.2	5.2	***
Total area cultivable land owned	3.7	4.5	***

Note: P-values = ***< 1%, **< 5%, *< 10%.

Select descriptive characteristics of household survey respondents in Arusha, according to their targeting intervention arm, are presented in Table 15. As mentioned in chapter 3, respondents in villages of treatment type PL and PV were chosen from among eligible households identified via a public meeting where the eligibility criteria were enforced. In contrast, eligible farmers in villages of treatment type VL and VV were provided by the VVC. As a result, if the VVC were to use different criteria to establish eligibility, one would observe differences in the characteristics across village types. This is not what we find in Table 3, however, as there is general exchangeability between households (all considered eligible by the respective VVCs) in each of the four intervention arms in Arusha.

Table 15: Select demographic and socioeconomic characteristics of households in Arusha in the baseline survey, according to targeting intervention arm

Characteristic	Targeting intervention arm (village)				F or χ^2 statistic
	PL	VL	PV	VV	
Female-headed households	36%	37%	41%	42%	0.731
Where household head has no formal education	18%	18%	21%	18%	0.922
Mean age of household head	49.6 years	49.4 years	47.3 years	49.0 years	0.731
Reporting no literacy (reading or writing)	18%	26%	22%	20%	0.532
Cultivable land area	2.1 ha	2.1 ha	2.2 ha	1.9 ha	0.733
Housing wealth index	0.4	0.6	0.3	0.3	0.330
Connectedness to position holders	14%	17%	11%	14%	0.661
Reporting severe food insecurity	24%	19%	23%	27%	0.561

5.2.1 Impact of NAIVS on input use

This section addresses evaluation questions:

1) What is the impact of NAIVS on:
    a) Production and welfare at the beneficiary level?
        (i) Direct returns: input use, productivity, profitability, quantity of output sold on the market.
    b) Heterogeneous treatment effects and spillovers within villages?
    c) How are the non-eligible or non-participating households affected by the programme?
    d) Does the introduction of participatory village meetings to identify eligible farmers and a lottery to select beneficiaries result in less efficient targeting?

The theory of change underlying NAIVS notes that the provision of the input subsidy is expected to increase the use of improved inputs and productivity among maize and rice farmers. Households are expected to produce more per acre as a result of the inputs and should, therefore, be able to sell, rather than solely consume, more of their crops and experience increased profitability from farming. Though only a limited number of households will have benefited from NAIVS, the demonstration effect of positive outcomes from the use of improved inputs is anticipated to encourage non-beneficiary farmers to purchase the inputs commercially. For non-farming households and/or those who do not produce enough for household consumption, there should be a benefit of reduced maize and rice prices over time as the local supply increases.

Given that NAIVS involves a subsidy for a productive input, there is a potential tension between efficiency and equity. If the eligibility rules were relaxed, the VVC could target the most productive individuals, thus enhancing efficiency and the potential for demonstration effects. However, the most productive individuals may also be those already purchasing inputs in the market and so the programme could crowd out investment. Relatedly, these beneficiaries could turn out to be better-off farmers, and so, while efficient, beneficiary selection would not be equitable.

If, on the other hand, the rules for eligibility were enforced, the yield gains may not materialise as less productive households may become beneficiaries. In this case, a secondary market could arise to restore efficiency. The least productive beneficiaries could sell the inputs to non-beneficiary productive farmers for a price between zero and the cost of the inputs in the market. The price will be determined by the number of productive farmers competing for the subsidised inputs.

This discussion suggests that the following predictions should be born out in the data:
(1) in PL and VL villages, beneficiary farmers should report having transferred, sold or given inputs to non-beneficiaries;
(2) as a result, the difference in input usage between beneficiaries and non-beneficiaries should be highest in PV and VV villages and lowest in PL and VL villages (recall that in PV and VV villages, beneficiaries were selected according to the standard VVC practices, whereas in PL and VL villages beneficiaries were selected according to a lottery); and
(3) higher input usage in VV villages should translate into higher yield gains as well.

Table 15 reports the effect of the programme on inputs transferred, purchased and used in plots devoted to maize.

The dependent variable in column 1 is a dummy variable that takes value 1 if the respondent reported a non-beneficiary friend to whom he or she had given agricultural inputs. The number of observations is the number of reported friends. Some respondents reported more than one friend, while others reported none. In total, 326 respondents reported 529 friends. If the lottery in villages of type PL and VL led allocated vouchers to less productive beneficiaries, one would expect more transfers of inputs from unproductive beneficiaries to (productive) non-beneficiaries. Table 15 also reports the p-value of a t-test of the difference in coefficient between beneficiaries and not in each village. The p-values in column 1 suggest that, in PV villages (just like VV villages), beneficiaries are not more likely to transfer inputs than non-beneficiaries. In fact, input transfers are not reported. In contrast, in VL villages, beneficiaries are more likely to report transfers than non-beneficiaries, as expected. PL villages look similar to VL villages, but the p-value is not significant at conventional levels (p-value is 0.179). Column 2 reports area planted for all plots devoted to maize. At follow-up there were 1,075 maize plots owned by 867 households. We find that programme beneficiaries did not cultivate larger plots, nor did they use more improved seeds in their plots (not reported). Columns 3 and 4 describe fertiliser use. Interestingly, beneficiaries in VV villages use around one more bag of DAP fertiliser (basal) than non-beneficiaries from those villages. This would be expected if beneficiaries were more productive, and so as we hypothesised above, the difference in input usage between beneficiaries and non-beneficiaries in VV (and PV) villages is the highest. In contrast, beneficiaries in PL and VL villages are significantly less likely to use DAP (and fertiliser in general). The p-values for PL and VL villages are above 0.100, suggesting no difference in fertiliser application between beneficiaries and non-beneficiaries, while in PV villages, the p-value is lower, suggesting a difference. Finally, column 5 reports the use of hired labour. Here the pattern is not as clear since beneficiaries in VL villages appear to be more likely to hire labour relative to non-beneficiaries. In PL villages, we again find no differences between beneficiaries and not.

5.2.2 Impact of NAIVS on agricultural productivity and profitability

This section addresses evaluation question:

1) What is the impact of NAIVS on:
    a) Production and welfare at the beneficiary level?
        i) Direct returns: input use, productivity, profitability, quantity of output sold on the market.

As discussed in the previous section, NAIVS is expected to increase crop production among beneficiaries, which will, in turn, increase the amount of household production sold and the profitability of smallholder agriculture. Table 16 reports several measures of crop productivity, including crop failure, yield, profitability and the percentage of production sold.

Crop Failure: Column 1 of Table 16 concerns crop failure. As it turns out, non-beneficiaries in PL and VL villages were less likely to suffer from complete crop failure, compared to those in VV (and PV) villages. The p-values suggest a similar story to that of Table 15, without differences between beneficiaries and non-beneficiaries in PL and VL villages, but with differences in PV villages. Given that non-beneficiaries in PL and VL villages may have enjoyed improved seeds and fertiliser, the use of these inputs may have led to better coping strategies in a drought situation like the one experienced in 2011–2012.

Table 16: Impact of the NAIVS programme on agricultural productivity—regression models predicting input usage, crop failure and yields among follow-up household respondents in Arusha, stratified by plot-level and household-level analyses

	Gives inputs to a non-beneficiary	Maize area planted (acres)	Kg of basal fertiliser per acre	Kg of all fertiliser per acre	Hired labour (1=Yes)
	OLS	OLS	Tobit	Tobit	OLS
	(1)	(2)	(3)	(4)	(5)
Beneficiary	-0.041	-0.769	40.062*	15.606	0.034
	(0.056)	(1.99)	(21.00)	(10.81)	(0.07)
PL village	-0.099*	-2.149	26.453	7.986	0.075
	(0.052)	(1.72)	(21.82)	(10.57)	(0.07)
VL village	-0.136***	0.297	0.194	1.284	-0.01
	(0.045)	(2.18)	(25.33)	(11.44)	(0.07)
PV village	-0.069	-0.121	-15.992	-1.673	-0.032
	(0.055)	(2.39)	(23.26)	(10.99)	(0.07)
Beneficiary x PL	0.095	0.593	-54.219*	-11.158	-0.047
	(0.069)	(1.99)	(27.68)	(13.77)	(0.10)
Beneficiary x VL	0.168**	-1.405	-20.909	-1.173	0.142
	(0.072)	(2.41)	(30.90)	(14.39)	(0.10)
Beneficiary x PV	0.031	-1.324	24.384	14.919	0.057
	(0.071)	(2.60)	(29.63)	(14.59)	(0.10)
Pct of censured observations			0.91	0.63	
R-squared	0.024	0.012			0.014
Mean of dep. var. in VV	0.113	3.5	6.2	16.6	0.411
P-val. of B = Non-B in…					
PL village	0.179	0.262	0.430	0.601	0.859
VL village	0.005	0.112	0.399	0.126	0.011
PV village	0.808	0.211	0.003	0.002	0.168
N. contacts	529				
N. plots		1,075	1,075	1,075	1,075
N. households	299	867	867	867	867

Note: The models presented here represent an intention to treat analysis and are not adjusted for any baseline covariates. Clustered SE at the individual level in parenthesis. Levels of significance = *p < 0.10, **p < 0.05, ***p < 0.01.

Yields: Column 2 reports yields per acre, treating crop failures as 0; given the censored structure of the data, a Tobit specification is employed. As we hypothesised, we find that beneficiaries in VV villages have higher yields. Further, we find that non-beneficiaries in PL and VL (and PV) villages also experience higher yields than those of VV villages. Though the VVC selected beneficiaries through their own system in PV villages, their pool from which to select was limited by the public eligibility meeting, and the PV village targeting was not as efficient as in VV villages. The p-values report again that in PL and VL villages, perhaps due to lower input use among beneficiaries and the corresponding higher input use among non-beneficiaries, there is no difference in yields between beneficiaries and not. In contrast, in PV and VV villages, beneficiaries have significantly higher yields than non-beneficiaries.

Thus, beneficiaries in PL and VL villages may be selling or giving away the inputs obtained through the voucher programme to some of the non-beneficiaries, consistent with column 1 in Table 16. In PV and VV villages, in contrast, beneficiaries may use all the inputs into their own plots. This might happen if VVCs target the most productive individuals when selecting beneficiaries in PV and VV villages, while in PL and VL villages, the lottery results in the selection of less productive beneficiaries and/or those less able to afford the top-up.

Percentage of output sold: Column 3 reports the percentage of total output sold in the market. Consistent with the theory, beneficiaries in villages of type PV and VV sell a higher proportion of their output to the market, while there is no difference for farmers living in villages of type PL.

Table 17: Impact of the NAIVS programme on agricultural productivity –– regression models predicting input usage, crop failure and yields among follow-up household respondents in Arusha, and heterogeneity in impacts when household head is female

	Crop failure	**Log total yield**	**Pct of output sold**
	Probit	**Tobit**	**OLS**
	(1)	(2)	(3)
Beneficiary	-0.208	0.804**	0.086**
	(0.209)	(0.397)	(0.044)
PL village	-0.446**	1.135***	0.083*
	(0.216)	(0.363)	(0.044)
VL village	-0.554***	0.994***	0.045
	(0.211)	(0.347)	(0.041)
PV village	-0.460**	0.813**	-0.011
	(0.210)	(0.362)	(0.037)
Beneficiary x PL	0.122	-0.645	-0.063
	(0.311)	(0.503)	(0.062)
Beneficiary x VL	0.184	-0.398	0.011
	(0.323)	(0.490)	(0.059)
Beneficiary x PV	-0.414	0.149	0.032
	(0.339)	(0.475)	(0.057)
Pct of censured observations		0.10	
R-squared			0.029
Mean of dep. var. in VV	0.178	4.5	0.187
P-val. of B = Non-B in…			
PL village	0.709	0.605	0.600
VL village	0.924	0.158	0.017
PV village	0.020	0.000	0.002
N. plots	1,075	1,075	
N. households	867	867	775

Note: The models presented here represent an intention to treat analysis and are not adjusted for any baseline covariates. Clustered SE at the individual level in parenthesis. Levels of significance = *p < 0.10, **p < 0.05, ***p < 0.01.

While these findings support the hypotheses underlying the targeting interventions, i.e., we see evidence of sharing inputs between beneficiaries and non-beneficiaries in all but VV villages and higher yields in VV villages, it is unclear what to conclude about the intended

impacts of NAIVS. When targeting criteria are enforced transparently, we do not find the intended impacts, but that may be because beneficiaries are not following the programme design by purchasing and applying inputs. It may be the case that beneficiaries selected by the VVC in VV villages are following the NAIVS protocol, which results in higher yields compared to non-beneficiaries; however, these farmers may have already been using improved inputs and are simply now purchasing them at a subsidised rate. This type of displacement of commercial input purchases in VV villages would do nothing for the macro-level impacts on productivity and food prices that NAIVS aims to achieve.

5.2.3 Impacts on input use and agricultural productivity by gender of household head

This section addresses evaluation question:

1) Are returns to fertiliser and input use heterogeneous by gender of the household head?

The gender analysis of the NAIVS study attempts to understand differential impacts of the project on men and women. While the project is anticipated to increase improved input use and overall productivity, it is not clear that women and men will experience similar levels of benefits. This could be due to varying areas of land under cultivation or varying ability to meet the labour demands on fertilised plots. Tables 18A, 18B, 19A and 19B present the effect of the programme on the same set of agriculture outcomes described above, stratifying results by the gender of the household head to explore heterogeneity in effects.

Area planted and input use: Focusing on the plot-level results (Tables 18A and 19A), male and female beneficiaries did not cultivate larger plots compared to male and female non-beneficiaries respectively. Columns 3 and 4 in each table describe fertiliser use. We find that in PV and VV villages, being a programme beneficiary is associated with a significant increase in fertiliser use among women (VV: β = 41.53, p < 0.001, PV: β = 13.24, p = 0.007) but not among men (VV: β = –7.393, PV: β = 13.26, p = 0.145). These findings may be due to the fact that in PV and VV villages, the VVC is preferentially selecting better women farmers who have either used fertiliser in the past or are more capable of adopting the new technology when given the opportunity through the voucher scheme. In contrast, the male farmers selected in VV villages may already be fertiliser users and thus being a beneficiary has a smaller effect. However, we do not see the same pattern in PL villages, where there is no difference between men and women non-beneficiaries, and beneficiary status is not associated with increased fertiliser use among women or men. In PV villages, female beneficiaries are significantly more likely than non-beneficiaries to use fertilizer.

Crop failure: Column 1 in Tables 18B and 19B concerns crop failure. There are no strong differences in crop failure between beneficiaries and non-beneficiaries for male- or female-headed households.

Yields: Column 2 in Tables 18B and 19B reports yields per acre, treating crop failures as 0; given the censored structure of the data, a Tobit specification is employed. In VL and PV villages, beneficiaries in female-headed households have significantly larger yields as compared to non-beneficiaries. No difference is observed between female beneficiaries and non-beneficiaries in VV villages, which is surprising given our hypothesis that the normal selection process implemented in VV villages should have produced beneficiaries that were better able to benefit from the input subsidy. Further, as described above, fertiliser use was indeed higher among female beneficiaries.

The pattern is slightly different among male-headed households, where beneficiaries in VV and PV villages have larger log total yields as compared to non-beneficiaries in the same village. The larger yield among beneficiaries in VV villages is not surprising given potential preferential selection of those households best prepared to benefit from the programme. The findings for PV villages (among both male- and female-headed households) are consistent with the increased input use in these villages described above.

In PL villages, there is no difference in yields by beneficiary status among either women or men, which is consistent with the observation that there was no increased adoption of inputs in PL villages. Thus, true randomised implementation of the intervention may have resulted in selection of farmers that were not able to benefit from the programme, resulting in increased selling or transfers of inputs in these villages.

Percentage of output sold: Column 3 in Tables 18B and 19B reports on the percentage of output sold. Female beneficiaries in VL and PV villages shared a larger percentage of the output as compared to female beneficiaries. The same was not observed among male-headed households (column 3, Table 19B), where the percentage of output sold among beneficiaries was greater only in PV villages.

5.2.4 Impact of NAIVS on household nutrition

This section addresses the following evaluation questions:

1) What is the impact of NAIVS on:
    a) Production and welfare at the beneficiary level?
       Indirect returns: nutrition, household welfare.

The expected long-term impacts of NAIVS, which stem from increased production and profitability from household farming, include improved nutrition and education. Greater crop production will support households in meeting dietary needs. In addition, the anticipated increase in household income from crop sales would allow households to purchase diverse and more nutritious foods, which would lead to improved nutritional outcomes.

Food insecurity: Food insecurity (FI) is the lack of physical, social and economic access to sufficient food for dietary needs and food preferences. Food insecurity was determined from the responses to a subset of three questions from the Household Food Insecurity Access Scale (HFIAS) (Coates, Swindale and Bilinsky 2007). Respondents were asked how often, in the last four weeks, they worried that their household would not have enough food, how often they were not able to eat preferred foods because of lack of resources and whether anyone in the household went to bed hungry. Based on these responses, households were classified into two mutually exclusive groups:
(1) food secure or moderate food insecurity; or
(2) severe food insecurity.
Severe food insecurity was defined as at least one household member going to bed hungry or 'often' worrying (more than 10 days in the last month) about food access or not being able to eat preferred foods.

Table 18A: Impact of the NAIVS programme on agricultural productivity—regression models predicting input usage, crop failure and yields among follow-up household respondents in Arusha, and heterogeneity in impacts when household head is male

	Gives inputs to a non-beneficiary	Area planted	Kg of basal fert. per acre	Kg of fertiliser per acre	Hired labour (1=Yes)
	OLS	OLS	Tobit	Tobit	OLS
	(1)	(2)	(3)	(4)	(5)
Beneficiary	-0.043	-1.447	80.519***	41.528***	0.159*
	(0.119)	(3.47)	(30.88)	(14.25)	(0.10)
PL village	-0.052	-3.027	38.461	28.616*	0.151
	(0.132)	(3.09)	(38.12)	(15.33)	(0.10)
VL village	-0.143	0.312	-60.800	-3.793	-0.038
	(0.097)	(3.79)	(44.65)	(15.31)	(0.10)
PV village	-0.143	-3.033	14.329	17.375	0.063
	(0.097)	(3.09)	(32.99)	(14.12)	(0.09)
Beneficiary x PL	0.008	1.310	-101.797**	-44.759**	-0.146
	(0.159)	(3.48)	(46.49)	(19.77)	(0.14)
Beneficiary x VL	0.126	-1.293	-8.038	1.286	0.151
	(0.145)	(4.13)	(54.53)	(20.12)	(0.14)
Beneficiary x PV	0.093	1.272	3.99	-4.133	0.017
	(0.129)	(3.48)	(41.43)	(19.71)	(0.13)
Pct of censured observations			0.88	0.75	
R-squared	0.034	0.021			0.043
Mean of dep. var. in VV	0.118	4.0	7.5	13.6	0.427
P-val. of B = Non-B in…					
PL village	0.738	0.494	0.530	0.812	0.903
VL village	0.314	0.220	0.127	0.003	0.002
PV village	0.323	0.525	0.002	0.007	0.057
N. contacts	126				
N. plots		513	513	513	513
N. households	205	415	415	415	415

Note: The models presented here represent an intention to treat analysis and are not adjusted for any baseline covariates. Clustered SE at the individual level in parenthesis. Levels of significance = *p < 0.10, **p < 0.05, ***p < 0.01.

Table 18 B: Impact of the NAIVS programme on agricultural productivity—regression models predicting input usage, crop failure and yields among follow-up household respondents in Arusha, and heterogeneity in impacts when household head is female

	Crop failure	Log total yield	Pct of output sold
	Probit	Tobit	OLS
	(1)	(2)	(3)
Beneficiary	-0.031	0.548	0.082
	(0.280)	(0.596)	(0.050)
PL village	-0.769**	1.596***	0.092*
	(0.340)	(0.483)	(0.056)
VL village	-0.548*	0.954**	0.020
	(0.282)	(0.485)	(0.047)
PV village	-0.284	0.722	0.017
	(0.267)	(0.518)	(0.044)
Beneficiary x PL	0.430	-0.938	-0.068
	(0.450)	(0.739)	(0.079)
Beneficiary x VL	-0.284	0.220	0.065
	(0.443)	(0.694)	(0.075)
Beneficiary x PV	-0.446	0.314	0.017
	(0.421)	(0.728)	(0.072)
Pct of censured observations		0.13	
R-squared			0.042
Mean of dep. var. in VV	0.218	4.2	0.145
P-val. of B = Non-B in…			
PL village	0.257	0.374	0.807
VL village	0.358	0.031	0.008
PV village	0.129	0.039	0.056
N. plots	513	513	
N. households	415	415	368

Note: The models presented here represent an intention to treat analysis and are not adjusted for any baseline covariates. Clustered SE at the individual level in parenthesis. Levels of significance: = *p < 0.10, **p < 0.05, ***p < 0.01.

Table 19A: Impact of the NAIVS programme on agricultural productivity—regression models predicting input usage, crop failure and yields among follow-up household respondents in Arusha, and heterogeneity in impacts when household head is male

	Gives inputs to a non-beneficiary OLS	Area planted OLS	Kg of basal fert. per acre Tobit	Kg of fertiliser per acre Tobit	Hired labour (1=Yes) OLS
	(1)	(2)	(3)	(4)	(5)
Beneficiary	-0.094	0.125	6.155	-7.393	-0.087
	(0.070)	(1.50)	(26.93)	(15.51)	(0.10)
PL village	-0.130**	-1.108	14.399	-10.609	0.004
	(0.066)	(0.78)	(26.06)	(14.53)	(0.10)
VL village	-0.158***	0.401	17.184	4.021	0.018
	(0.060)	(1.67)	(30.32)	(15.72)	(0.10)
PV village	-0.029	3.147	-40.693	-14.66	-0.138
	(0.085)	(3.49)	(32.19)	(16.44)	(0.10)
Beneficiary x PL	0.194**	-0.301	-11.998	18.969	0.039
	(0.088)	(1.52)	(34.00)	(18.97)	(0.14)
Beneficiary x VL	0.222**	-1.490	-8.061	-0.381	0.137
	(0.088)	(2.11)	(36.75)	(19.59)	(0.14)
Beneficiary x PV	0.016	-4.325	36.986	27.915	0.123
	(0.099)	(3.74)	(42.60)	(20.99)	(0.14)
Pct of censured observations			0.87	0.72	
R-squared	0.048	0.021			0.018
Mean of dep. var. in VV	0.125	3.0	4.9	19.5	0.395
P-val. of B = Non-B in…					
PL village	0.699	0.438	0.778	0.288	0.617
VL village	0.140	0.361	0.939	0.517	0.595
PV village	0.154	0.220	0.196	0.145	0.705
N. contacts	324				
N. plots		565	565	565	565
N. households	173	452	452	452	452

Note: The models presented here represent an intention to treat analysis and are not adjusted for any baseline covariates. Clustered SE at the individual level in parenthesis. Levels of significance = *p < 0.10, **p < 0.05, ***p < 0.01.

Table 19B: Impact of the NAIVS programme on agricultural productivity—regression models predicting input usage, crop failure and yields among follow-up household respondents in Arusha, and heterogeneity in impacts when household head is male

	Crop failure	Log total yield	Pct of output sold
	Probit	Tobit	OLS
	(1)	(2)	(3)
Beneficiary	-0.386	0.940*	0.075
	(0.318)	(0.522)	(0.071)
PL village	-0.239	0.711	0.059
	(0.299)	(0.519)	(0.068)
VL village	-0.546*	1.030**	0.064
	(0.316)	(0.484)	(0.068)
PV village	-0.765**	0.897*	-0.047
	(0.349)	(0.494)	(0.060)
Beneficiary x PL	-0.124	-0.323	-0.038
	(0.468)	(0.662)	(0.094)
Beneficiary x VL	0.561	-0.913	-0.033
	(0.462)	(0.666)	(0.092)
Beneficiary x PV	.	-0.012	0.060
	.	(0.601)	(0.089)
Pct of censured observations		0.08	
R-squared			0.024
Mean of dep. var. in VV	0.137	4.9	0.228
P-val. of B = Non-B in…			
PL village	0.139	0.130	0.549
VL village	0.602	0.948	0.482
PV village	0.225	0.002	0.013
N. plots	489	565	
N. households	395	452	407

Note: The models presented here represent an intention to treat analysis and are not adjusted for any baseline covariates. Clustered SE at the individual level in parenthesis. Levels of significance = *p < 0.10, **p < 0.05, ***p < 0.01.

As Table 20 highlights, nearly one quarter (23 per cent) of the 2012 Arusha sample reported severe food insecurity at their follow-up interview. There were no significant differences across villages overall (p=0.33). In bivariate analyses, severe food insecurity was significantly higher among non-beneficiaries as compared to beneficiaries (26 versus 21 per cent, p = 0.03). Female-headed households were also significantly more likely to report severe food insecurity (29 versus 18 per cent, p < 0.001). In particular, women in PL villages were more likely to be severely food insecure compared to men in PL villages (p = 0.06). This same pattern was observed in VV villages (p < 0.01).

Table 20: Percentage of households with severe food insecurity in Arusha follow-up survey, stratified by beneficiary status and gender of household head

Characteristic	Overall N (%)	Village			
		PL	VL	PV	VV
All households	215 (23)	56 (23)	51 (21)	47 (21)	61 (28)
Beneficiary status:					
Beneficiaries	98 (21)	26 (20)	22 (18)	20 (18)	30 (26)
Non-beneficiaries	117 (26)	30 (27)	29 (25)	27 (25)	31 (30)
Gender of household head:					
Female	128 (29)	32 (29)	28 (25)	28 (26)	40 (36)
Male	87 (18)	24 (19)	23 (18)	19 (17)	21 (19)

In order to estimate the impact of being an NAIVS beneficiary in the 2010–2011 season on food security in 2012, we estimated three linear probability models (LPMs). The first examined the impact of NAIVS by targeting the intervention arm (village PL, VL, PV or VV), while the second and third were stratified by the gender of the household head. The full models are included in Table 13. Here, we present results in particular related to the PL and VV villages in Arusha, whereby vouchers were assigned by a public meeting to determine eligibility and a lottery to determine beneficiaries, or were allocated using standard procedures (VVC) respectively. As noted in Table 15, food security status was balanced at baseline across the four village types. In PL villages, where we have the cleanest estimate of impact, receipt of the voucher is associated with a 6.5-point reduction in the probability of being severely food insecure, although this was non-significant (p = 0.241). In VV villages, although the direction of the effect is similar, it is smaller and, again, not statistically significant (β = –0.034, p = 0.572) (Table 21, column 1). When we repeat the analysis using logistic regression to account for the binary outcome, our interpretation remains the same: there is no strong effect on severe food insecurity. For example, in PL villages, being an NAIVS beneficiary is associated with a 30 per cent reduction in the odds of being severely food insecure compared to non-beneficiaries (p = 0.238).

We then examine gender-disaggregated effects of the programme on severe food insecurity; results from this analysis are presented in columns 2 and 3 of Table 21. Like in the overall models, we do not see any effect of the programme among men or women, except we see a small, marginally significant effect among women in VL villages. In these villages, 18 per cent of female beneficiaries were severely food insecure compared to 32 per cent of female non-beneficiaries (p = 0.085).

Table 21: Impact of the NAIVS programme on severe household food insecurity—regression models predicting severe food insecurity in Arusha, and heterogeneity in impacts by gender of the household head

	Severe food insecurity	Male-headed households: severe food insecurity	Female-headed households: severe food insecurity	Severe food insecurity
	OLS	OLS	OLS	Logistic
	(1)	(2)	(3)	(4)
Beneficiary	-0.0344	-0.0625	0.0128	-0.171
	(0.0607)	(0.0781)	(0.0914)	(0.302)
PL village	-0.0274	0.00160	-0.0317	-0.135
	(0.0614)	(0.0807)	(0.0937)	(0.302)
VL village	-0.0474	-0.0537	-0.0342	-0.240
	(0.0601)	(0.0795)	(0.0880)	(0.303)
PV village	-0.0452	-0.0216	-0.0600	-0.229
	(0.0612)	(0.0831)	(0.0888)	(0.309)
Beneficiary x PL	-0.0304	-0.0276	-0.0663	-0.190
	(0.0821)	(0.104)	(0.127)	(0.430)
Beneficiary x VL	-0.0346	0.0676	-0.153	-0.243
	(0.0806)	(0.104)	(0.122)	(0.438)
Beneficiary x PV	-0.0371	-0.00945	-0.0773	-0.256
	(0.0822)	(0.106)	(0.125)	(0.449)
Constant	0.295***	0.229***	0.351***	-0.870***
	(0.0447)	(0.0612)	(0.0638)	(0.214)
Obs.	920	478	442	920
Adj. R-squared	0.001	-0.007	0.001	
P-val. of B = Non-B in…				
PL villages	0.241	0.189	0.545	0.238
VL villages	0.194	0.940	0.085	0.193
PV villages	0.198	0.318	0.447	0.198

Note: The models presented here represent an intention to treat analysis and are not adjusted for any baseline covariates. SE in parentheses. Levels of significance = *p < 0.10, **p < 0.05, ***p < 0.001.

Dietary diversity: Dietary diversity is an alternate way to measure food insecurity and is defined as the number of unique food groups consumed over a given period of time. Reduced dietary diversity is often used as a household coping strategy in response to stresses, consistent with the predictable patterns of loss management that households employ to alleviate the deleterious effects of shocks. For example, when facing a real or potential shock, households typically stabilise their consumption of staples (grains or tubers) and reduce consumption of more nutrient dense foods like eggs, vegetables, meat and diary (Block *et al.* 2007; Lesotho DMA 2008). The World Health Organization and UNICEF's framework for childhood nutrition posits that increasing access to and production of food may have effects on childhood nutrition (Bellamy 1998; Smith and Haddad 2000). In this regard, we hypothesised that increased production or income resulting from the NAIVS programme may result in household diet diversification, especially in settings such as rural Tanzania, where there is an overreliance on starchy staple foods and very low consumption of nutrient-rich foods.

Dietary diversity was assessed only during the follow-up (2012) survey using the Household Dietary Diversity Score (HDDS), a proxy measure for the nutritional quality of a household's diet (Swindale and Bilinsky 2006).

The HDDS ranges from 0 to 12 and is the number of different food groups consumed in the 24 hours preceding the interview, out of the following 12 groups: (1) cereals; (2) roots and tubers; (3) pulses, legumes and nuts; (4) vegetables; (5) fruits; (6) meat and poultry; (7) eggs; (8) fish and seafood; (9) milk and milk products; (10) oils and fats; (11) sugar and sweets; and (12) condiments and miscellaneous. As with food insecurity, we focus our analysis on the PL and VV villages.

Overall, the mean dietary diversity score among households in Arusha in the follow-up survey was 6.3 (SD 2.1.9). Female-headed households reported a significantly lower mean dietary diversity score as compared to male-headed households (6.0 versus 6.6, $p < 0.01$). However, there was no difference in dietary diversity score according to the household's beneficiary status (ignoring the targeting village); the mean dietary diversity score was 6.3 in both groups ($p=0.71$) (Table 22).

Table 22: Mean dietary diversity in households in Arusha follow-up survey, stratified by beneficiary status and gender of household head

Characteristic	Mean	Village			
		PL	VL	PV	VV
All households	6.3	6.4	6.4	6.2	6.3
Beneficiary status:					
Beneficiaries	6.3	6.2	6.5	6.4	6.3
Non-beneficiaries	6.3	6.5	6.2	6.1	6.4
Gender of household head:					
Female	6.0	6.0	6.1	6.1	5.9
Male	6.6	6.7	6.6	6.4	6.8

In order to estimate the impact of NAIVS on dietary diversity, we ran three OLS regression models with dietary diversity score as the outcome. In the first model including all households (column 1, Table 23), we find that there is no beneficial effect of being an NAIVS beneficiary on dietary diversity in any of the villages. In subsequent models (columns 2 and 3 of Table 23), we examine the effects of the programme disaggregated by gender. There is consistently no positive effect of the programme on dietary diversity. We first consider the effect on men (Model 2). Among male-headed households in PL villages, the mean dietary diversity among beneficiaries was 6.6 compared to 6.9 among non-beneficiaries. Among male-headed households in VV villages, the mean dietary diversity among beneficiaries was 6.6 compared to 7.0 among non-beneficiaries. Considering the effect on women, among those in PL villages, the mean dietary diversity among female beneficiaries was 5.9 compared to 6.0 among non-beneficiaries. Among female-headed households in VV villages, the mean dietary diversity among beneficiaries was 5.9 compared to 5.8 among non-beneficiaries.

There are several possible explanations for these null findings. First, the sample sizes are small after stratifying by several dimensions, including gender; thus, we likely have very low power to detect any effects. Second, it is possible that the NAIVS programme had no effect on the quality of household diets, even if severe food insecurity was moderately decreased (which does not seem likely). Third, it is possible that perhaps the programme increased consumption of cereals such as maize but did not increase consumption of other nutrient-dense foods such

as meats, fish, or fruits and vegetables. In this case, there could be increased caloric consumption but no change in dietary diversity. This scale would not detect such an increase in energy intake.

Table 23: Impact of the NAIVS programme on household dietary diversity—regression models predicting dietary diversity in Arusha, and heterogeneity in impacts by gender of the household head

	Household Dietary Diversity Score OLS	Male-headed households: Dietary Diversity Score OLS	Female-headed households: Dietary Diversity Score OLS
	(1)	(2)	(3)
Beneficiary	-0.0923	-0.346	0.0306
	(0.290)	(0.456)	(0.342)
PL village	0.183	-0.0561	0.179
	(0.273)	(0.406)	(0.348)
VL village	-0.157	-0.628	0.225
	(0.282)	(0.437)	(0.352)
PV village	-0.297	-0.753*	0.0670
	(0.278)	(0.432)	(0.344)
Beneficiary x PL	-0.218	-0.0147	-0.146
	(0.379)	(0.571)	(0.476)
Beneficiary x VL	0.383	0.745	0.0400
	(0.375)	(0.575)	(0.464)
Beneficiary x PV	0.394	0.611	0.287
	(0.376)	(0.583)	(0.461)
Constant	6.362***	6.979**	5.842***
	(0.214)	(0.337)	(0.258)
Obs.	920	478	442
Adj. R-squared	-0.002	-0.001	-0.010
P-val. of B =Non-B in…			
PL villages	0.204	0.295	0.729
VL villages	0.220	0.254	0.822
PV villages	0.207	0.465	0.304

Note: The models presented here represent an intention to treat analysis and are not adjusted for any baseline covariates. SE in parentheses. Levels of significance = *p < 0.10, **p < 0.05, ***p < 0.001.

5.2.5 Impact of NAIVS on Household Welfare and Human Capital Accumulation

This section addresses the following evaluation questions:

1) What is the impact of NAIVS on:
    a) Production and welfare at the beneficiary level.
        i) Indirect returns: household welfare, human capital accumulation, women's autonomy.

NAIVS is expected to increase profitability and income for beneficiary households over time as a result of higher yields per plot. Assuming households have more disposable income from the increased sales of crops, the theory of change hypothesises that these resources would be channelled towards school fees or other investments, such as adult labour, which would facilitate school enrolment and attendance. Women may also experience an additional indirect return of the project in the form of greater autonomy over assets and decisionmaking in the household as a result of greater earnings. The study examines household expenditures to investigate the effect on incomes and spending habits and school enrolment to understand whether increased incomes translate to better educational outcomes. It also investigates the question of impacts on women's empowerment.

Household expenditures: We examined the effect of the NAIVS programme on monthly per capita household expenditures. Monthly per capita expenditures in Tanzanian shillings (TZS) was computed from reported expenditures on food and transportation in the last four weeks and on non-food household items, education and medical expenses in the last 12 months. The average monthly per capita expenditure was TZS36,510 (about USD22). This differed significantly by beneficiary status in the unadjusted bivariate analysis (TZS32,684 among non-beneficiaries versus TZS40,055 among beneficiaries, $p < 0.01$), but it did not differ by gender of the household head (TZS35,073 among male-headed households versus TZS38,067 among female-headed households, $p = 0.281$).

The results of the regressions by village type are presented in Table 24 with Ln (per capita expenditures) as the dependent variable. The first model examines the effect of the programme in the overall sample and finds that there is no effect of the programme on per capita expenditures in PL, VL or VV villages. However, we see a significant effect in beneficiary villages: the average monthly per capita expenditure is TZS24,929 in non-beneficiary households compared to TZS37,033 in beneficiary households ($p = 0.046$). When we examine the model stratified by gender of the household head (columns 2 and 3 in Table 24), we see that the effect is mainly among women in VL villages. Here, the vouchers were allocated by the public meeting and VVC selection (not by lottery). This effect may explain the effect we note earlier on severe food insecurity among female beneficiaries in VL villages.

Table 24: Impact of the NAIVS programme on household expenditures—regression models predicting monthly per capita expenditures in Arusha, and heterogeneity in impacts by gender of the household head

	Ln(monthly per capita expenditures) OLS	Male-headed households: Ln(monthly per capita expenditures) OLS	Female-headed households: Ln(monthly per capita expenditures) OLS
	(1)	(2)	(3)
Beneficiary	0.102	0.192	0.0132
	(0.111)	(0.159)	(0.154)
PL village	0.147	0.0613	0.290*
	(0.106)	(0.144)	(0.157)
VL village	0.158	0.251	0.0740
	(0.107)	(0.160)	(0.142)
PV village	0.0649	0.0688	0.0664
	(0.108)	(0.157)	(0.150)
Beneficiary x PL	-0.0811	-0.0526	-0.161
	(0.147)	(0.203)	(0.214)
Beneficiary x VL	0.102	-0.140	0.383*
	(0.151)	(0.215)	(0.214)
Beneficiary x PV	0.0245	-0.106	0.164
	(0.147)	(0.208)	(0.210)
Constant	10.01***	9.970***	10.05***
	(0.0816)	(0.116)	(0.115)
Obs.	919	478	441
Adj. R-squared	0.009	0.001	0.020
P-val. of B = Non-B in…			
PL villages	0.830	0.269	0.323
VL villages	0.046	0.717	0.008
PV villages	0.194	0.520	0.217

Note: The models presented here represent an intention to treat analysis and are not adjusted for any baseline covariates. SE in parentheses. Levels of significance = *p < 0.10, **p < 0.05, ***p < 0.001.

Proportion of household children enrolled in school: We hypothesised that if participation in NAIVS resulted in increased income for households, some of that increased income might be directed towards school expenses for children in the household, given that poverty strongly predicts school attendance in this region. Therefore, we examined the effect of NAIVS on the proportion of household children ages 7 to 18 currently enrolled in school. Table 25 summarises the proportion of children in school according to village type, beneficiary status and gender of the household head. Overall, school enrolment was high, as households on average reported that 82 per cent of children ages 7 to 18 were currently in school. Although the proportion of children in school was qualitatively different between beneficiaries and non-beneficiaries (83 per cent versus 81 per cent), this difference did not achieve statistical

significance (p = 0.210). Similarly, although female-headed households reported a slightly larger proportion of household children being enrolled in school as compared to male-headed households (83 per cent versus 81 per cent), this difference was also not statistically significant (p = 0.189).

Table 25: Proportion of children ages 7 to 18 in school among households in Arusha follow-up survey, stratified by beneficiary status and gender of household head

Characteristic	Mean	Village			
		PL	VL	PV	VV
All households	0.822	0.816	0.827	0.814	0.828
Beneficiary status:					
Beneficiaries	0.830	0.811	0.822	0.846	0.845
Non-beneficiaries	0.813	0.821	0.833	0.781	0.810
Gender of household head:					
Female	0.830	0.843	0.821	0.822	0.836
Male	0.812	0.790	0.833	0.803	0.820

There may be differences in enrolment because the voucher programme may induce income effects and since education is a normal good, enrolment rates may change. For example, if vouchers in VVC villages go to well-to-do households, enrolment rates may not change as much as in lottery villages, where the vouchers may be assigned to poorer households who may respond more to the increased income from the vouchers. In order to estimate the impact of NAIVS on the proportion of children in school by village type, we ran three OLS regression models with the proportion of children in school as the outcome. These regressions are presented in column 1, Table 26. Overall, the regressions do not offer any evidence of a differential effect of investments in education by beneficiary status within each village. Interestingly, beneficiary status results in an *increased* proportion of children in school only in VV villages (β = 0.031). In both PL and VL villages, beneficiaries report a reduced proportion of children in school (β = –0.036 and -0.026 respectively) and in PV villages there is essentially no difference between beneficiaries and non-beneficiaries (β = –0.001). However, none of these results achieve statistical significance.

Recognising that households may divert resources differently to girl and boy children in the household, we also explored whether participation in NAIVS influenced the proportion of girls ages 7 to 18 currently enrolled in school. The results of these regressions (among households overall, and then separately among male- and female-headed households) are presented in Table 26. Similar to the findings among children overall, there was no statistically significant effect of beneficiary status on the proportion of girls in school.

Table 26: Impact of the NAIVS programme on children's enrolment in school—regression models predicting the proportion of household children currently enrolled in school in Arusha, and heterogeneity in impacts by gender of the household head

	Proportion of children in school OLS	Male-headed households: proportion of children in school OLS	Female-headed households: proportion of children in school OLS
	(1)	(2)	(3)
Beneficiary	0.0366	0.0448	0.0306
	(0.84)	(0.70)	(0.52)
PL village	0.0108	0.00840	0.0194
	(0.25)	(0.13)	(0.32)
VL village	0.0231	0.0351	0.0129
	(0.54)	(0.54)	(0.23)
PV village	-0.0293	-0.0460	-0.0154
	(-0.66)	(-0.67)	(-0.26)
Beneficiary x PL	-0.0473	-0.0773	-0.0268
	(-0.79)	(-0.88)	(-0.32)
Beneficiary x VL	-0.0498	-0.0472	-0.0582
	(-0.83)	(-0.54)	(-0.70)
Beneficiary x PV	0.0283	0.0567	0.00347
	(0.45)	(0.60)	(0.04)
Constant	0.810***	0.798***	0.821***
	(26.15)	(16.87)	(20.04)
Obs.	727	364	363
Adj. R-squared	-0.005	-0.008	-0.016
P value of B = Non-B in…			
PL village	0.384	0.262	0.899
VL village	0.525	0.838	0.456
PV village	0.982	0.871	0.842

Note: The models presented here represent an intention to treat analysis and are not adjusted for any baseline covariates. SE in parentheses. Levels of significance = *p < 0.10, **p < 0.05, ***p < 0.001.

Table 27: Impact of the NAIVS programme on girls' enrolment in school—regression models predicting the proportion of female household children currently enrolled in school in Arusha, and heterogeneity in impacts by gender of the household head

	Proportion of girls in school OLS (1)	Male-headed households: proportion of girls in school OLS (2)	Female-headed households: proportion of girls in school OLS (3)
Beneficiary	0.0222	-0.0102	0.0536
	(0.46)	(-0.15)	(0.79)
PL village	0.0242	0.00324	0.0524
	(0.50)	(0.05)	(0.74)
VL village	-0.0121	0.0385	-0.0614
	(-0.25)	(0.55)	(-0.91)
PV village	0.0178	-0.0294	0.0545
	(0.36)	(-0.39)	(0.81)
Beneficiary x PL	-0.0472	-0.0221	-0.0769
	(-0.69)	(-0.23)	(-0.78)
Beneficiary x VL	-0.0189	-0.0289	-0.0150
	(-0.28)	(-0.30)	(-0.16)
Beneficiary x PV	-0.0411	0.0362	-0.107
	(-0.57)	(0.35)	(-1.08)
Constant	0.855***	0.860***	0.851***
	(24.78)	(17.28)	(17.69)
N	568	294	274
Adj. R-squared	-0.010	-0.020	-0.006
P value of B = Non-B in…			
PL village	0.632	0.783	0.719
VL village	0.513	0.884	0.274
PV village	0.651	0.926	0.466

Note: The models presented here represent an intention to treat analysis and are not adjusted for any baseline covariates. SE in parentheses. Levels of significance = *p < 0.10, **p < 0.05, ***p < 0.001.

Women's autonomy: We hypothesised that participation in the NAIVS programme may have an impact on women's autonomy in female-headed households as a result of potential increased earnings or reduced food insecurity. Women's autonomy was assessed using the Household Decision-Making Scale (Nanda 2011). This scale collapses women's responses to three questions about who usually makes decisions about major household purchases, daily household needs, and visits to family and relatives. Women who indicated that someone else primarily made these decisions scored 0 for each question, while women who indicated that they made these decisions alone or jointly with a spouse or partner scored 1, resulting in an overall score ranging from 0 to 3. Overall, most respondents from female-headed households (N = 442) described having high decision-making authority (mean score of 2.74, SD 0.74). Only 5 per cent of the sample indicated that they could not make any decisions related to major household purchases, daily household purchases or how to spend their free time. The majority (87 per cent) indicated that they could make decisions about the three alone or jointly with a spouse or partner. In bivariate analyses, there was no statistically significant difference in household decision-making authority between beneficiaries and non-beneficiaries (scores of 2.72 and 2.77 respectively, results not shown).

In an OLS regression model analysing the impact of being a beneficiary on women's autonomy score by village (Table 28), compared to women in VV villages, women in PL, VL and PV villages reported higher but not statistically significant differences in autonomy. In particular, women in PL villages scored approximately 0.3 points higher on the autonomy scale as compared to women in VV villages, and this difference was marginally significant ($p = 0.063$). However, beneficiary status was not associated with increased autonomy, except for women in VV villages. Again, these results did not reach statistical significance.

Table 28: OLS regression predicting household decision-making autonomy among female heads of household

	β coefficient (SE)
Beneficiary	0.192 (0.153)
PL village	0.278* (0.149)
VL village	0.15 (0.162)
PV village	0.233 (0.151)
PL x beneficiary	-0.398* (0.205)
VL x beneficiary	-0.302 (0.218)
PV x beneficiary	-0.276 (0.197)
Constant	2.6
R-squared	0.0138
N	422

Note: The models presented here represent an intention to treat analysis and are not adjusted for any baseline covariates. SE in parentheses. Levels of significance = *p < 0.05.

6. Conclusions

Over the past two years, an impact evaluation has been implemented to determine the impact of NAIVS on the agricultural productivity, health and nutrition, and overall welfare of farming households across Tanzania. Data have been collected from multiple sources, including farming households participating in NAIVS and those eligible but not participating in the programme, local agro-dealers involved in implementing NAIVS and village leadership responsible for selecting NAIVS beneficiaries. In data collection activities, female-headed farming households, who face numerous constraints to productive agricultural livelihoods yet comprise a significant portion of agricultural labour, have been oversampled in order to examine heterogeneous impacts of the programme by gender. In addition, a qualitative study focused on female farmers was conducted to explore in-depth the unique barriers and facilitators to NAIVS participation among this population.

Analyses of these data reveals mixed findings regarding the impact of NAIVS on agricultural production, household nutrition and household welfare. In our analysis of the impacts of the subsidy programme on agricultural input use, productivity and profitability in one region (Arusha) where targeting interventions were successful, we have identified the tension between efficiency and equity. In villages where the local VVC chose beneficiaries, farmers who were most able to benefit from the programme appear to have been targeted more efficiently

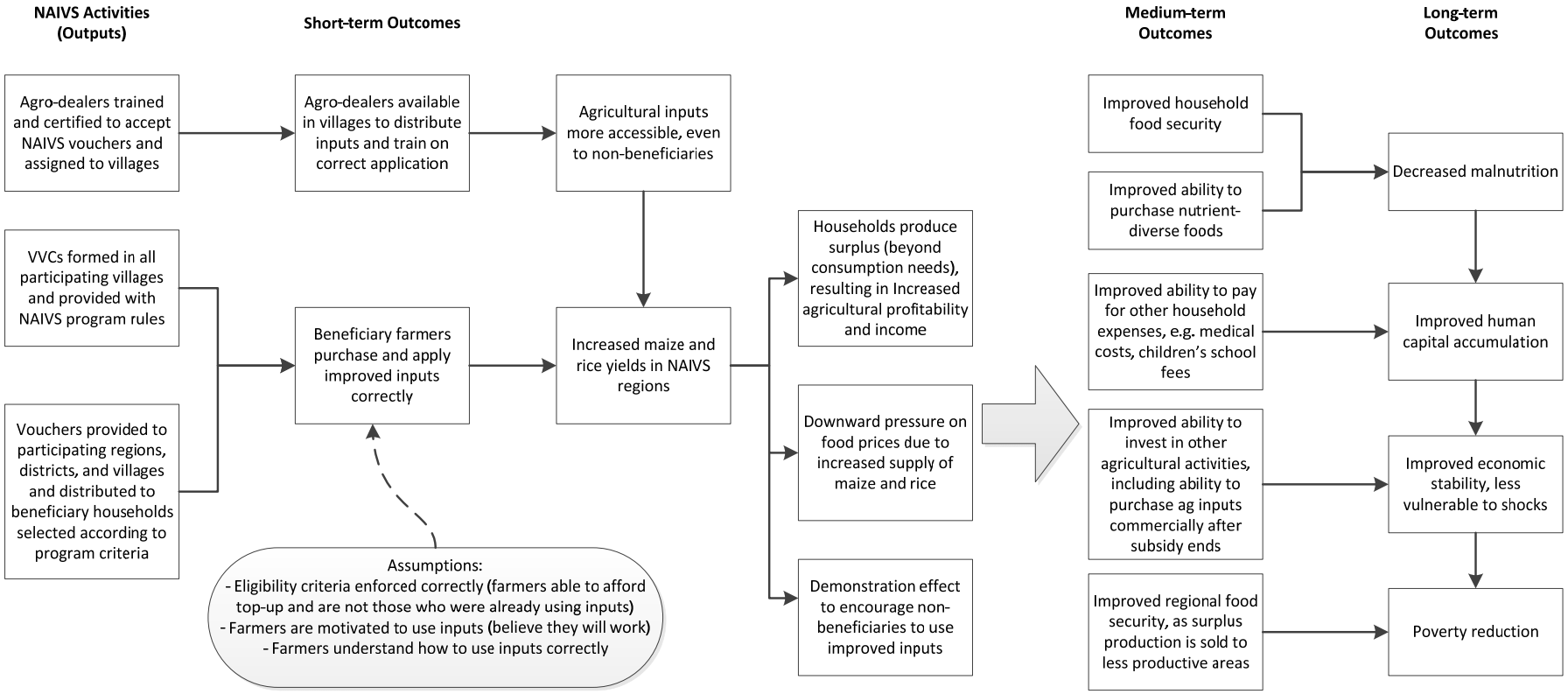
than in other villages in which beneficiaries were randomly chosen, resulting in higher use of inputs and in some cases yields among households in these villages. Our findings also suggest that in villages where farmers were randomly selected to receive the voucher, selling or sharing of vouchers between beneficiaries and non-beneficiaries was more common, thus highlighting spillover effects of the programme. In order to enhance efficiency while providing for the most in need, the eligibility criteria should be strictly enforced while allowing the transfer and sale of vouchers or its contents, something that the current programme implementers vehemently oppose. Assuming the better-performing farmers targeted by the VVC were not already using improved inputs, which would mean NAIVS simply resulted in a displacement of commercial purchases and little overall adoption, or yield or production increases, this recommendation would allow NAIVS to achieve its intended outcomes.

Though NAIVS intended to target mid-income smallholder farmers, the question of how to reach poorer households who still may be able to benefit from a small productive input is worth further consideration. A significant proportion of these poorer households are female-headed households, which the programme aims to prioritise, but they remain disproportionately underrepresented as beneficiaries. Our findings clearly highlighted the disenfranchised position of female-headed households relative to their male counterparts. In the baseline household and listing surveys, female-headed households reported having fewer assets, owning less land and living in poorer housing construction. Further, food insecurity and reduced dietary diversity were significantly more common among female-headed households. Results from the qualitative study confirmed and contextualised these findings, revealing that although female farmers felt positively about the programme and its ability to help boost yields, many could not afford the top-up payment and, as a result, did not participate. However, our impact analyses also demonstrated that being a beneficiary was associated with increases in input use, yields and output sold among female-headed households in all villages except those where voucher assignment was fully randomised. This likely represents a combination of preferential selection of female farmers who were able to benefit from the programme, but also meant larger gains for female-headed households when compared to their male counterparts. Smaller input packages or flexible payment terms could make participation in the programme more affordable for poorer households and the resulting marginal returns to this demographic in terms of production, food security and incomes will likely be higher than average.

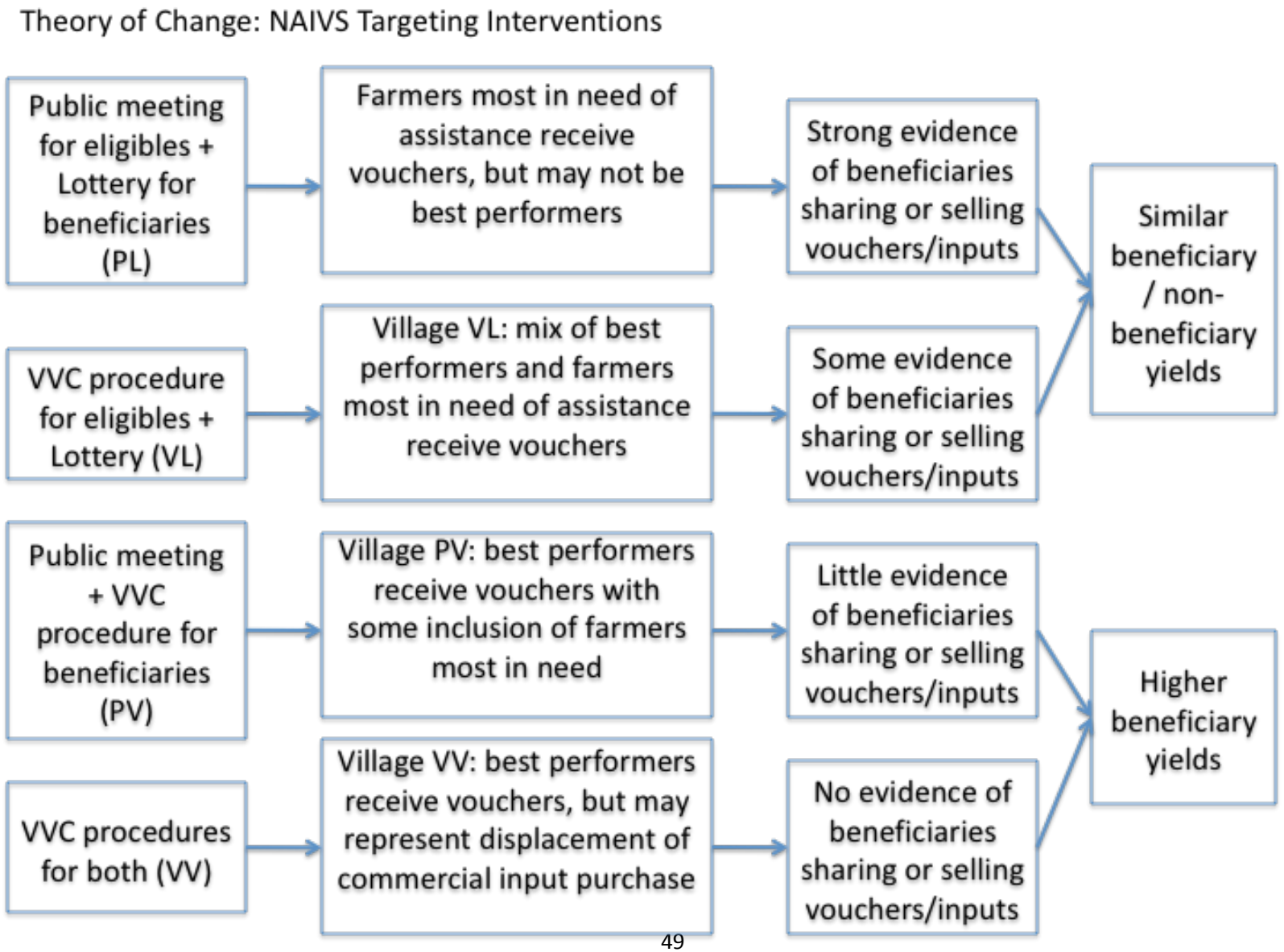
Although we hypothesised that participation in NAIVS might influence other distal outcomes related to household expenditures, human capital accumulation and autonomy (among female-headed households), our impact analyses did not consistently detect significant differences between beneficiaries and non-beneficiaries with regard to these outcomes. We only measured outcomes after one year of participation in NAIVS, and it is possible that these differences might emerge only after several years of sustained participation. Therefore, these results should not be taken as conclusive evidence that agricultural subsidy programmes cannot influence health, welfare and other outcomes, perhaps only that they do not do so in the short term.

Appendix A: Theory of change

NAIVS Theory of Change



Appendix B: Theory of change



References

Alderman, Harold, 2002. Do local officials know something we don't? Decentralization of targeted transfers in Albania. *Journal of Public Economics*, 83(3), pp.375–404.

Baird, Sarah, McIntosh, Craig and Özler, Berk, 2011. The regressive demands of demand-driven development. *Policy Research Working Paper Series 5883*. Washington, DC: World Bank, 2011.

Bellamy, C, 1998. *The state of the world's children*. Geneva: United Nation's Children's Fund (UNICEF).

Block, SA, Kiess, L, Webb, P, Kosen, S, Moench-Pfanner, R, Bloem, MW and Timmer CP, 2004. Macro shocks and micro outcomes: child nutrition during Indonesia's crisis. *Economics & Human Biology*, 2(1), pp.21–44.

Chirwa, Ephraim, Matita, Mirriam and Dorward, Andrew, 2002. Targeting agricultural input subsidy coupons in Malawi. London: Centre for Development, Environment and Policy, School of Oriental and African Studies (SOAS), University of London, and Wadonda Consult (unpublished).

Coates, J, Swindale, A and Bilinsky P, 2007. *Household Food Insecurity Access Scale (HFIAS) for measurement of food access: indicator guide.* Washington, DC: United States Agency for International Development (USAID).

Cogill, B, 2003. Anthropometric indicators measurement guide. Food and Nutrition Technical Assistance Project. Washington, DC: Academy for Educational Development.

Duflo, E, Kremer, M and Robinson, J (forthcoming). Nudging farmers to use fertilizer: theory and experimental evidence from Kenya. *American Economic Review*.

Duflo, E, Kremer, M and Robinson, J, 2008. How high are rates of return to fertilizer? Evidence from field experiments in Kenya. *American Economic Review Papers and Proceedings*, 98(2), pp.482–488.

Economic Commission for Africa, 2006. Mitigating the impact of HIV/AIDS on smallholder agriculture, food security, and rural livelihood in Southern Africa: challenges and action plan. Johannesburg: United Nations Economic Commission for Africa (UNECA).

Feder, G, Just, R and Zilberman D, 1985. Adoption of agricultural innovations in developing countries: a survey. *Economic Development and Cultural Change*, 33(2), p.255–298.
Fernald, L, Gertler, P and Neufeld, L, 2008. Ten-year impact of Oportunidades: Mexico's conditional cash transfer program on child growth, cognition, language and behavior. *Lancet*, 374(9706), pp.1997–2005.

Gillespie, S and Drimie, S, 2009. Hyperendemic AIDS, food insecurity, and vulnerability in southern Africa: a conceptual evolution working paper. Washington, DC: International Food Policy Research Institute and RENEWAL.
Gillespie, S and Kadiyala, S, 2005. HIV/AIDS and food and nutrition security: from evidence to action. Washington, DC: International Food Policy Research Institute.

Jack, Kelsey, 2013. Private information and the allocation of land use subsidies in Malawi. *American Economic Journal: Applied Economics*, 5(3), pp.113–135.
Joint World Health Organization (WHO)/Centers for Disease Control and Prevention (CDC), 2004. Technical Consultation on the Assessment of Iron Status at the Population Level. Geneva: WHO/CDC.

Kim, J, Pronyk, P, Barnett, T and Watts, C, 2008. Exploring the role of economic empowerment in HIV prevention. *Aids*, 22(Suppl. 4), pp.S57–71.

Lesotho Disaster Management Authority (DMA), Lesotho Vulnerability Assessment Committee (LVAC), United Nations World Food Programme 2008. Vulnerability and food insecurity in urban areas of Lesotho: An assessment of the high prices on vulnerable households on ten major cities.

Nanda, G, 2011. *Compendium of gender scales.* Washington, DC: FHI 360/C-Change.
National Bureau of Statistics (NBS), 2007–2008. Tanzania HIV/AIDS and malaria indicator survey: preliminary report. Dar es Salaam: NBS.

Quisumbing, A and Pandolfelli, L, 2008. *Promising approaches to address the needs of poor female farmers*. Washington, DC: International Food Policy Research Institute.
Rowling, M, 2008. Women farmers toil to expand Africa's food supply. Reuters, 26 December.

Smith, L and Haddad, L, 2000. *Overcoming child malnutrition in developing countries: past achievements and future choices*. Washington, DC: International Food Policy Research Institute.

Suri, T (forthcoming). Selection and comparative advantage in technology adoption. *Econometrica*.

Swindale, A and Bilinsky, P, 2006. Household Dietary Diversity Score (HDDS) for measurement of household food access: indicator guide (v.2). Washington, DC: Food and Nutrition Technical Assistance Project, Academy for Educational Development.
UNAIDS, 2010. *Report on the global AIDS epidemic*. Geneva: UNAIDS.

United Republic of Tanzania, 2008. *The Economic Survey 2007*. Dar es Salaam: United Republic of Tanzania.

World Bank. 2009. *Gender in agriculture sourcebook*. Washington, DC: World Bank.

World Health Organization (WHO), 2004. *Indicators for assessing vitamin A deficiency and their application in monitoring and evaluating intervention programmes*. Geneva: WHO.

World Health Organization, World Food Program, United Nations Standing Committee on Nutrition and United Nations Children's Fund, 2007. Community-based management of malnutrition. A joint statement. Geneva: World Health Organization, World Food Program, the United Nations Standing Committee on Nutrition and the United Nations Children's Fund.

UNAIDS, 2010. *Report on the global AIDS epidemic*. Geneva: UNAIDS.

Publications in the 3ie Impact Evaluation Report Series

The following reports are available from http://www.3ieimpact.org/evidence-hub/impact-evaluation-repository/

The promise of preschool in Africa: A randomised impact evaluation of early childhood development in rural Mozambique, *3ie Impact Evaluation Report 1*. Martinez, S, Naudeau, S and Pereira, V (2012)

A rapid assessment randomised-controlled trial of improved cookstoves in rural Ghana, *3ie Impact Evaluation Report 2*. Burwen, J and Levine, DI (2012)

The GoBifo project evaluation report: Assessing the impacts of community-driven development in Sierra Leone, *3ie Impact Evaluation Report 3*. Casey, K, Glennerster, R and Miguel, E (2013)

Does marginal cost pricing of electricity affect groundwater pumping behaviour of farmers? Evidence from India, 3ie Impact Evaluation Report 4. Meenakshi, JV, Banerji, A, Mukherji, A and Gupta, A (2013)

Impact evaluation of the non-contributory social pension programme 70 y más in Mexico, *3ie Impact Evaluation Report 5.* Rodríguez, A, Espinoza, B, Tamayo, K, Pereda, P, Góngora, V, Tagliaferro, G and Solís, M (2014)

The impact of daycare on maternal labour supply and child development in Mexico, *3ie Impact Evaluation Report 6*. Angeles, G, Gadsden, P, Galiani, S, Gertler, P, Herrera, A, Kariger, P and Seira, E (2014)

Social and economic impacts of Tuungane: final report on the effects of a community-driven reconstruction programme in the Democratic Republic of Congo, *3ie Impact Evaluation Report 7*. Humphreys, M, Sanchez de la Sierra, R and van der Windt, P (2013)

Paying for performance in China's battle against anaemia, *3ie Impact Evaluation Report 8*. Zhang, L, Rozelle, S and Shi, Y (2013)

No margin, no mission? Evaluating the role of incentives in the distribution of public goods in Zambia, *3ie Impact Evaluation Report 9*. Ashraf, N, Bandiera, O and Jack, K (2013)

Truth-telling by third-party audits and the response of polluting firms: Experimental evidence from India, 3ie Impact Evaluation Report 10. Duflo, E, Greenstone, M, Pande, R and Ryan, N (2013)

An impact evaluation of information disclosure on elected representatives' performance: evidence from rural and urban India, 3ie Impact Evaluation Report 11. Banerjee, A, Duflo, E, Imbert, C, Pande, R, Walton, M and Mahapatra, B (2014)

Targeting the poor: evidence from a field experiment in Indonesia, 3ie Impact Evaluation Report 12. Atlas, V, Banerjee, A, Hanna, R, Olken, B, Wai-poi, M and Purnamasari, R (2014)

Scaling up male circumcision service provision: results from a randomised evaluation in Malawi, 3ie Impact Evaluation Report 13. Thornton, R, Chinkhumba, J, Godlonton, S and Pierotti, R (2014)

Providing collateral and improving product market access for smallholder farmers: a randomised evaluation of inventory credit in Sierra Leone, 3ie Impact Evaluation Report 14. Casaburi, L, Glennerster, R, Suri, T and Kamara, S (2014)

A youth wage subsidy experiment for South Africa, 3ie Impact Evaluation Report 15. Levinsohn, J, Rankin, N, Roberts, G and Schöer, V (2014)

The impact of mother literacy and participation programmes on child learning: evidence from a randomised evaluation in India, 3ie Impact Evaluation Report 16. Banerji, R, Berry, J and Shortland, M (2014)

Assessing long-term impacts of conditional cash transfers on children and young adults in rural Nicaragua, *3ie Impact Evaluation Report 17*. Barham, T, Macours, K, Maluccio, JA, Regalia, F, Aguilera, V and Moncada, ME (2014)

Impact of malaria control and enhanced literacy instruction on educational outcomes among school children in Kenya: a multi-sectoral, prospective, randomised evaluation, 3ie Impact Evaluation Report 18. Brooker, S and Halliday, K (2014)

A randomised evaluation of the effects of an agricultural insurance programme on rural households' behaviour: evidence from China, 3ie Impact Evaluation Report 19. Cai, J, de Janvry, A and Sadoulet, E (2014)

Environmental and socioeconomic impacts of Mexico's payments for ecosystem services programme, *3ie Impact Evaluation Report 20*. Alix-Garcia, J, Aronson, G, Radeloff, V, Ramirez-Reyes, C, Shapiro, E, Sims, K and Yañez-Pagans, P (2014)

Shelter from the storm: upgrading housing infrastructure in Latin American slums, *3ie Impact Evaluation Report 21*. Galiani, S, Gertler, P, Cooper, R, Martinez, S, Ross, A and Undurraga, R (2014)

A wide angle view of learning: evaluation of the CCE and LEP programmes in Haryana, 3ie Impact Evaluation Report 22. Duflo, E, Berry, J, Mukerji, S and Shotland, M (2014)

Enhancing food production and food security through improved inputs: an evaluation of Tanzania's National Agricultural Input Voucher Scheme with a focus on gender impacts, *3ie Impact Evaluation Report 23*. Gine, Xavier, Patel, Shreena, Cuellar-Martinez, Cristina, McCoy, Sandi and Ralph Lauren (2015)

In Tanzania, agriculture accounts for 27 per cent of GDP, 80 per cent of employment and 75 per cent of household income. It is a key component of the country strategy for poverty reduction. The National Agriculture Input Voucher Scheme (NAIVS), implemented since 2008 by the Ministry of Agriculture, Food Security and Cooperatives, provides a 50 per cent subsidy, through vouchers, for the purchase of chemical fertiliser and improved seed to maize and rice farmers. This study evaluates the impact of NAIVS across Tanzania.

The findings do not show impacts on incomes, food security or educational attainment in the short term. However, the disenfranchised position of female-headed households relative to their male counterparts is clearly indicated. Female farmers were positively inclined toward the programme and its ability to help boost agricultural yields. But many of them could not afford the top-up payment and, as a result, did not participate. To make NAVIS more effective, the study recommends that the programme should be targeted better and a provision for transfer and sale of vouchers included.

Impact Evaluation Series

International Initiative for Impact Evaluation
202-203, Rectangle One
D-4, Saket District Centre
New Delhi – 110017
India

3ie@3ieimpact.org
Tel: +91 11 4989 4444

www.3ieimpact.org