Does Subsidizing Legume Seeds Improve Farm Productivity and Nutrition in Malawi?

Makaiko G. Khonje, Christone Nyondo, Julius H. Mangisoni, Jacob Ricker-Gilbert, William J. Burke, William Chadza & Milu Muyanga

MwAPATA Institute
Working Paper No. 21/06

September 2021

P.O. Box 30883 Capital City, Lilongwe Malawi

Chilanga Drive, Off Blantyre Street, Area 10/44

Dr. Khonje (m.khonje@mwapata.mw) and Dr. Nyondo are Research Fellows at the MwAPATA Institute. Dr. Burke is a Research Director at the MwAPATA Institute. Dr. Mangisoni is a Professor of Agricultural Economics at the Lilongwe University of Agriculture and Natural Resources. Dr. Ricker-Gilbert is an Associate Professor of Agricultural Economics at Purdue University. Mr. Chadza is the Executive Director of the MwAPATA Institute. Dr. Muyanga is a Co-Principal Investigator of the MwAPATA Institute and Associate Professor in the Department of Agricultural, Food and Resource Economics at Michigan State University.

Early View:

Kindly note, this is an "Early View" version of this Working Paper. This means the analysis is considered ready to share, but the document has not undergone final copy editing. This paper can be cited or reproduced, but please check for updates to obtain the final version.

Does Subsidizing Legume Seeds Improve Farm Productivity and Nutrition in Malawi?

Makaiko G. Khonje, Christone Nyondo, Julius H. Mangisoni, Jacob Ricker-Gilbert, William J. Burke, William Chadza & Milu Muyanga

Executive Summary

Over the last two decades, most African governments have been implementing agricultural input subsidy programs (ISPs) aimed at increasing crop yields, incomes, and reducing hunger, nutritional insecurity and poverty. Although ISPs are popular policy interventions, it remains unknown whether ISPs improve productivity of nutrient-dense crops such as legumes, dietary quality, and child nutrition. We address this important gap by testing the hypothesis that subsidizing legume seeds improves farm productivity, dietary quality, and child nutrition. We use a decade-long nationally representative panel data from Malawi and panel regression models with instrumental variable approach to address potential endogeneity issues.

We find that subsidizing legume seeds increases area planted with legume crops, overall gross value of production, production and dietary diversity, calories, and micronutrient—vitamin A and zinc—consumption. We further found that subsidizing legume seeds is positively correlated with child weight-for-age Z-score but not height-for-age Z-score.

These novel findings emphasize that subsidizing legume seeds could be a valuable policy option to address malnutrition in the Malawian small farm sector.

1. Introduction

Hunger and different forms of malnutrition remain widespread public health problems in many developing countries, with sub-Saharan Africa (SSA) accounting for 40% of the global prevalence (Kinyoki et al., 2020; FAO et al., 2019; Unicef et al., 2019). While many factors contribute to different forms of malnutrition, a key factor is insufficient intake of nutrient-dense foods or too much intake of calorie-dense foods (Zaharia et al., 2021; Khonje et al., 2020). Compounded with declining soil fertility and low use of modern agricultural inputs, such as improved seeds and inorganic fertilizers, the challenge to feed the current and future generations with nutritious food is huge (Ariga et al., 2019; Snapp et al., 2014; Tittonell & Giller, 2013).

In response to this challenge, many countries are implementing agricultural input subsidy programs (ISPs), mainly for inorganic fertilizer and improved seeds for cereals like maize and rice. Emerging research (e.g., Abman & Carney, 2020; Ricker-Gilbert & Jayne, 2017; Mason & Ricker-Gilbert, 2013; Awotide et al., 2013) suggests that these input subsidies have contributed to increased cereal yields. However, in recent years, yields for cereals have typically remained low (below 3 tons/ha) compared to potential yields of over 5 tons/ha (Benson, 2021; FAOstat, 2021).

There are growing concerns that most ISPs are largely focusing on cereals and non-food crops like cotton and tobacco (e.g., Theriault & Smale, 2021; Wossen et al., 2017; Awotide et al., 2013), but not nutrient-dense crops such as legumes. Yet, legumes are key source of nutritious food for many smallholder farmers compared to cereals and animal-sourced foods (ASFs). ASFs are highly under consumed as they have high price relative to legumes, making them unattainable as part of a staple diet for many households (Hirvonen et al., 2020). Moreover, legumes may help to increase crop yields through biological enhancement of soil fertility (Snapp et al., 2014) and income (Rubyogo et al., 2019).

Though still under investigated, subsidizing legume seeds (SLSs) is one of the potential policy interventions that could improve household nutrition through productivity and income pathways. Even in countries where the ISP package includes legume seeds, so far, it remains unclear whether or not such subsidies can improve household nutrition through productivity and income mechanisms (Walls et al., 2018). For example, to what extent does SLSs increase area planted with legume crops, farm production diversity and productivity—measured in returns to land and labor—in smallholder farm households over time? Do members of households who acquired subsidized legume seeds have better diet quality and nutrition than non-beneficiaries?

To fill this important gap, the present article analyzes the effects of including legume seeds in an ISP on farm productivity and nutrition. We use a decade-long (2010-2020) nationally representative panel data from Malawi. Malawi is an excellent case study because its ISP included legume seeds, as the input subsidy program for maize, which has been implemented since 2004 and legume seeds were included as part of the subsidy package for farmers in various years of the program (Benson, 2021; Chirwa & Dorward, 2013). Moreover, child stunting rate is very high in Malawi (37%), and it is among the highest in the SSA region (Development Initiatives, 2020; Unicef et al., 2019).

Our contributions to the existing literature on farm input subsidies and nutrition is threefold. First, we provide new evidence on the effects of SLSs on farm productivity, dietary quality, and child nutrition. Though not on legume seed subsidies, a few studies (Smale et al., 2020; Harou, 2018) have found that fertilizer subsidies improved dietary quality. Other studies (Theriault & Smale, 2021; Chibwana et al., 2012) have found that input subsidies crowd out farmland allocated to legumes. Moreover, most previous studies (e.g., Theriault & Smale, 2021; Abman & Carney, 2020; Ricker-Gilbert & Jayne, 2017; Mason & Smale, 2013) have analyzed the effects of input subsidies on crop species diversity, maize yields, income, and poverty, but not on dietary quality and child nutrition.

Second, our empirical analyses use a decade-long nationally representative panel data from Malawi`s Integrated Household Panel Survey (IHPS), where smallholder farming households were interviewed four times between 2010 and 2020. A few related studies (Theriault & Smale, 2021; Smale et al., 2020; Chibwana et al., 2012) used datasets that were cross-sectional; where controlling for possible unobserved confounding factors is difficult. These previous studies used small samples while ours is nationally representative. In addition, to deal with potential endogeneity that could be caused by selection bias, measurement errors and omitted variable bias, we use panel (Mundlak) regression and instrumental variable (the control function) approach estimators. With these features, our results are likely to be externally valid and better accounting for unobserved heterogeneity.

Third, we contribute to the policy debate on how farm input subsidies (legume seed subsidies) influence nutritional outcomes such as dietary diversity, micronutrient consumption, and child nutrition in SSA. Lessons drawn from our study may provide practical policy interventions that could help to inform current or future ISPs as one of the potential pathways to addressing malnutrition in all its forms.

We found that SLSs increased area planted with legume crops, gross value of production, farm production and consumption (dietary) diversity, and micronutrient—vitamin A and zinc—consumption. We also found that SLSs is positively correlated with child weight-for-age Z-score (WAZ) but not height-for-age Z-score (HAZ). Our results suggest that inclusion of legume seeds in ISPs could improve nutrition through productivity and income pathways.

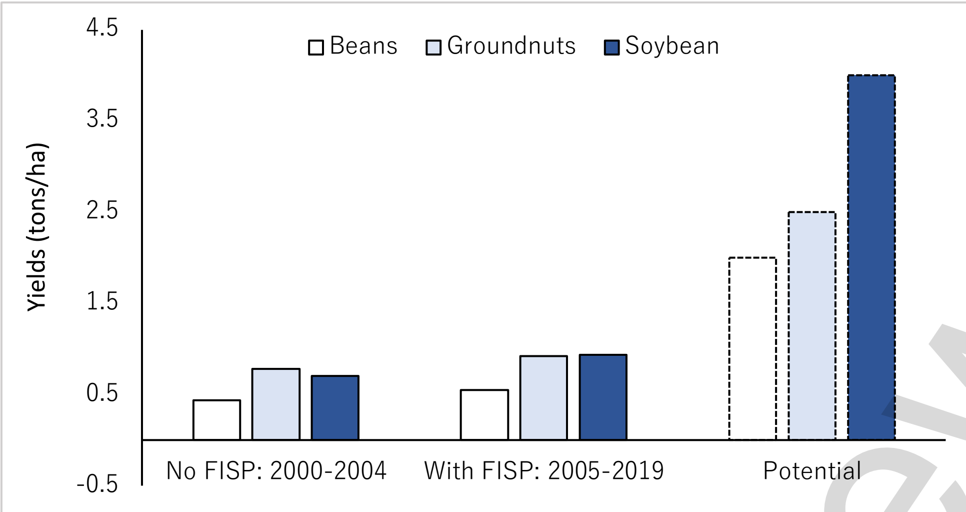
The rest of this article is organized as follows. In the following section, we provide background on ISP including legume seed in Malawi. We then present data and empirical strategy. The sections that follow present and discuss the results. The last section contains conclusions and implications for policy.

2. Legume seed subsidies in Malawi

Legumes are important for attainment of food and nutritional security and improving farm incomes for many smallholder farmers in SSA (Hirvonen et al., 2020; Rubyogo et al., 2019). However, achieving this goal, is often limited by low yields for legume crops. For example, yields for key legume crops have remained low (Figure 1) compared to potential yields of over 3 tons/ha. Possibly, three factors explain this observation. First, adoption rates for improved legume seeds are still low. Second, recycling of improved seeds beyond their vitality period is a common practice by most smallholder farmers. Third, due to limitations to access adequate fertilizer and traditional beliefs, inorganic fertilizers are rarely applied to legume crops.

To accelerate diffusion of improved legume seeds and improve crop yields, food and nutritional security and income, Malawi started implementing a large farm input subsidy program (FISP) in 2004/05 growing season (Benson, 2021; Chirwa & Dorward, 2013). The program (FISP), targeted about 0.9 to 1.5 million smallholder farmers, who could access subsidized inputs at 64-93% of the market price from 2004/05 to 2019/20 growing season

Figure 1: Yields for Legume Crops between 2000 and 2019 in Malawi (Average Yields)



Source: Authors' calculations. Data source: FAOstat; http://www.fao.org/faostat/en/#data/QC and Benson (2021).

(Abman & Carney, 2020; Harou, 2018; Chirwa & Dorward, 2013). In addition to two 50 kg bags of inorganic fertilizer—one basal, one urea—and improved—either hybrid or open pollinated variety—maize seeds (2-5 kg), farmers could buy legume seeds (1-2 kg) at reduced price using a coupon system (Benson, 2021; Abman & Carney, 2020; Holden & Lunduka, 2012).

Table 1 presents the quantity of subsidized inputs supplied to smallholder farmers in Malawi from 2010/11 to 2019/2020. We largely focus on the years that are used in our empirical analysis. In general, Table 1 suggests that the quantity of subsidized inputs—fertilizer, maize and legume seeds—distributed to smallholders in Malawi has declined significantly since 2010/11.

Unlike maize seed, the quantities of subsidized legume seeds have been relatively smaller over time (Table 1). By extension, the percentage of households using subsidized legume seeds in our sample, drastically declined over time than those who used commercial legume seeds (Figure 2). This suggests that there is substantial under investments or other structural barriers in the legume seed sector.

Table 1: Subsidized Input Supply by Survey Year in Malawi

Subsidized inputs supplied (MT)	2010/11 (1)	2013/14 (2)	2016/17 (3)	2019/20 (4)	Total (5)
Fertilizer	159,952	149,971	89,511	89,880	489,314
Maize seed	10,650	8,268	4,628	4,472	28,018
Legume seeds	2,727	3,042	1,664	1,283	8,715

Source: Authors' calculations. MT, metric tonnes. The data were drawn from Government of Malawi`s final report on the implementation of Agricultural Inputs Subsidy Program for 2010/11, 2013/14, 2016/17 and 2019/20 growing seasons. Disaggregated quantities are shown in Table A1 of the appendix.

One consequence of underinvestment in legume seed sector is that farmers are not maximizing yields from FISP due to several challenges (Lunduka et al., 2013). First, very few farmers were targeted in a village, which led to farmers sharing the inputs. Second, input suppliers were buying coupons from farmers and later claiming payments. Third, contracts were given to inefficient suppliers.

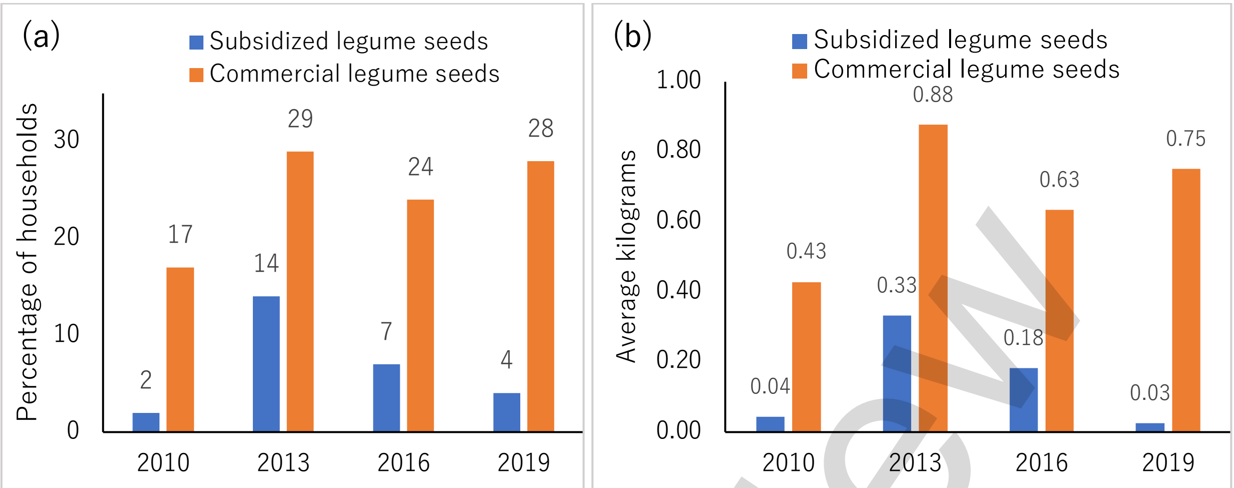
To address some of the aforementioned challenges from the past ISPs—i.e., FISP, the Malawi Government is implementing reforms under the Affordable Inputs Program (AIP), FISP's successor. Under the AIP program, almost all (3.6 million) smallholder farmers are accessing subsidized inorganic fertilizers and cereal (maize, rice and sorghum) seeds at 24% and 33% of the market price, respectively, using a biometric system (Chilundu, 2020). However, farmers are no longer accessing improved legume seeds. This could be a significant omission as most households in low-income countries get their proteins through legumes and not ASFs.

3. Data

Our empirical analysis uses a decade-long nationally representative panel data from Malawi`s IHPS. Smallholder farming households were interviewed four times—in 2010/2011, 2013, 2016/2017 and 2019/2020—over a decade. The IHPS datasets were collected by National Statistical Office (NSO) in collaboration with the World Bank.

A sample of 3246 households from 204 enumeration areas; representative for all districts in Malawi, were randomly selected for interviews in 2010. These households were re-interviewed every 3 years up to 2019. We use a total sample of 7034 smallholder farming households (1287, 1595, 1813 and 2339 households in 2010, 2013, 2016 and 2019,

Figure 2: Legume Seeds Used by Sampled Households, by Survey Year and Source



Source: Authors' calculations. (a) Percentage of households using legume seeds by survey year and source. (b) Average kilograms of legume seeds used by households, by survey year and source. N= 1287, 1595, 1813 and 2339 for 2010, 2013, 2016, and 2019, respectively.

respectively) in our analysis. The number of households increases overtime due to splitting of original households over the years. However, some households were lost due to death and attrition. Fortunately, the attrition rates were relatively low, about 3.8%, 5% and 4.5% in 2013, 2016 and 2019, respectively.

In all panel survey rounds, comprehensive information on the household composition, non-food consumption expenditure, asset ownership, and other socioeconomic variables were captured. These surveys also contain detailed data on agricultural production, food consumption and health outcomes including child anthropometric measures. We use most of these modules to generate selected variables of interest.

3.1 Measuring productivity and income

In this study, we used area planted with legume crops, farm production diversity (PD) and gross value of production (VoP) as indicators of farm production diversity and productivity or income, respectively. PD is calculated as the number of crops grown and livestock species produced by a farm household over the past 12 months. VoP (MK/ha) for all crops planted by the farming household is calculated as the ratio of value of crop output—price[1] (MK/kg) multiplied with quantity produced (kg)—to unit area (ha). Major crops grown in the study country are maize, groundnuts, common beans, sorghum, tobacco and soybean (see Table A2 in the appendix). Price data were drawn from Malawi`s Ministry of Agriculture.

While a better measure of productivity is crop yield (kg/ha), we opted to use gross value of production; because it is easy to deal with additive challenges. These two (PD and VoP) indicators have also been used elsewhere (e.g., Muthini et al., 2020; Muyanga & Jayne, 2019; Ecker, 2018; Sibhatu et al., 2015). However, as with most productivity outcomes, VoP may suffer from measurement errors especially if farmers use recall data to capture farm size (Wollburg et al., 2021).

3.2 Measuring dietary diversity and micronutrient consumption

Using household level (a 7-day dietary recall) food consumption data, we calculated two measures of consumption (dietary) diversity: the household dietary diversity score (HDDS) and food consumption score (FCS). HDDS is measured as the count of 12 different food groups (see panel A of Table A3 in the appendix for detailed food groups and their weights) consumed by the household during the recall period. HDDS has been widely used as an indicator of dietary diversity in several developing countries (e.g., Ecker, 2018; Sibhatu et al., 2015). Moreover, with HDDS, it is easy to measure dietary quality without serious measurement errors compared to calculating actual food/micronutrient consumption or intake (Villa et al., 2011). However, it has its own limitation, as some of the food groups included to calculate the indicator are unhealthy, especially sugars, sweets, and soft drinks (Muthini et al., 2020).

We also used FCS as a measure for dietary diversity. FCS is measured by summing a predetermined set of weights (see panel B of Table A3) designed to reflect the heterogeneous dietary quality for each of the 8 food groups consumed by the households (WFP, 2008). The FCS ranges in value from 0 to 112 and a higher score would imply a better heterogeneous dietary quality. FCS has been used as a measure of dietary quality in other

[1] The output (harvest) is valued using average price for April to December for each survey year. We used annual average prices to convert crop quantities produced by the farming household to VoP. To make VoP comparable across survey years, we used real output prices.

African countries (e.g., Hoddinott et al., 2018). Nevertheless, these two consumption diversity indicators (HDDS and FCS) do not measure the actual amount of food or (micro)nutrients consumed by the household.

To capture the multidimensional nature of nutrition, we also use household-level food consumption data and local food composition tables to calculate calories (Kcal/day/AE) and micronutrient—e.g., vitamin A (μg RE/day/AE), iron (mg/day/AE) and zinc (mg/day/AE)—consumption. These indicators are common proxy measures for dietary quality or nutrition. We adjusted calories and micronutrient consumption using male adult equivalents (AE). The use of household-level food consumption data to assess dietary quality has become common in the food economics literature (e.g., Ogutu et al., 2019; Zezza et al., 2017).

However, this approach does not account for the intra-household allocation of calories and micronutrients. With our available data, this could not be analyzed because the 24-hour food intake data were not captured during the surveys. Further, we acknowledge that the human body needs multiple micronutrients for a healthy life. However, deficiencies in vitamin A, iron, and zinc are relatively common in most developing countries and responsible for most health problems, especially among children (Zaharia et al., 2021; FAO et al., 2019; Unicef et al., 2019; Development Initiatives, 2018).

4. Empirical strategy

We hypothesize that accessing subsidized legume seeds could influence farm productivity and nutrition through various ways. First, legume seed subsidies may increase production, especially where farmers allocate more land to legume crops. Second, promotion of improved legume seeds under ISPs may help to increase crop yields through use of hybrid seeds and biological enhancement of soil fertility (Ricker-Gilbert & Jayne, 2017; Snapp et al., 2014). Third, the direct consumption of legumes—e.g., beans bio-fortified with vitamin A, iron, and zinc (Ogutu et al., 2020)—could improve dietary quality and nutritional outcomes of household members. Fourth, income from legume sales (Rubyogo et al., 2019) could be used to buy other nutritious foods as well as invest in other crops and livestock species. However, with the available data, it is not possible to conclusively model the income effect.

Here, we add to a few existing studies by testing the highlighted causal pathways; which have not been explicitly analyzed before. We use the following specification to analyze the effects of SLSs on farm productivity and nutrition:

$$\boldsymbol{y}_{ht} = \alpha + \beta \boldsymbol{SLS}_{ht} + \gamma_1 X_{ht} + \gamma_2 D_{ht} + \varepsilon_{ht} \tag{1}$$

where $\boldsymbol{y}_{ht}$ is the productivity outcome—e.g., area planted with legume crops, PD and gross value of production and dietary quality—e.g., HDDS, FCS, calories and micronutrient (vitamin A, iron and zinc) consumption—indicators of interest for household h at time t. The main explanatory variable of interest is $\boldsymbol{SLS}_{ht}$ and it is a binary variable that indicates whether a household received subsidized legume seeds or not. In addition, $\boldsymbol{SLS}_{ht}$ is measured as a continuous variable; quantities of subsidized legume seeds received by the household. The parameter of interest β measures whether accessing subsidized legume seeds improved farm productivity, dietary quality and child nutrition or not. We consider several household (X_{ht}) characteristics as controls in our regressions. D_{ht} represents a vector of time fixed effects. ε_{ht} is a random error term.

To address potential endogeneity bias from several sources, we use several empirical and identification strategies. First, our identification strategy relies on the use of household-level panel data that provide us with the ability to use the Mundlak approach—i.e., time varying variables are included as explanatory variables in panel estimators (Mundlak, 1978). The Mundlak estimator has two advantages. First, it removes correlation between the observed covariates and the time-constant unobservable factors that affect dependent variables of the interest (Wooldridge, 2019; Ricker-Gilbert & Chamberlin, 2018). Second, it is more efficient than the regular fixed effects (FE) estimator when the within variation in the data is smaller than the between variation (Debela et al., 2021).

We also use different control variables, which should be able to account for much of the correlation between observed covariates and unobserved time-varying shocks that could potentially bias our estimates. Lastly, we use Mundlak with a control function (CF) approach—we use district-level population shares overtime (Table A4 in the appendix) as an instrumental variable (IV)—to account for remaining potential endogeneity issues especially for non-linear models such as Poisson. District population shares for each survey year were drawn from Malawi`s NSO. Our results suggest that the IV is valid, because it is strongly correlated with the number of households who received subsidized legume seeds but uncorrelated with productivity and dietary quality outcomes (Table A5 in the appendix). This IV has also been used in similar studies (e.g., Abman & Carney, 2020). Nevertheless, with observational data addressing all sources of endogeneity may still be difficult.

We use different panel estimators—e.g., random effects (RE) and Mundlak—through Poisson and linear regression models based on characteristics of the dependent variable. For dependent variables that are measured with count data such as PD and HDDS, we use Poisson regression models for estimation. For dependent variables that are continuous and normally distributed such as area planted with legume crops, FCS, gross value of production, calories and micronutrient consumption, we use panel linear regression models.

5. Results and discussion

5.1 Descriptive results

Descriptive statistics for key variables of interest by survey year and access to subsidized legume seeds are shown in Tables 2, A6 and A7 of the appendix, respectively. On average, we find that 0.56 acres (equivalent to 43% of total land area) were planted with legume crops. As expected, we find that SLSs is positively correlated with area planted with legume crops (Table A7 of the appendix).

Over time, we find that for almost all productivity and dietary quality indicators, they follow an inverse U-shaped relationship. For example, the average consumption for vitamin A in 2010 was 693 and it increased to 1109 in 2013, and thereafter it declined to 750 in 2016 (Table 2). This may be associated with favorable weather conditions in 2012/2013 growing season where production estimates were marginal higher than in 2010 and 2016 (see Figure A1 in the appendix). Higher production levels for cereals and legumes may help smallholder farmers to improve diets through consumption of nutritious food from own production and market purchases.

However, interestingly, we further find that smallholder farmers who had received subsidized legume seeds have better diets and more income than non-beneficiaries over time. For example, on average, households who accessed subsidized legume seeds have higher HDDS and micronutrient consumption—e.g., vitamin A, iron and zinc—and gross VoP

Table 2: Descriptive Results for Key Variables by Survey Year

	All years	2010	2013	2016	2019
	(1)	(2)	(3)	(4)	(5)
Dependent variables					
Area planted with legumes (acres)	0.56	0.35	0.64	0.45	0.75
	(0.70)	(0.55)	(0.80)	(0.64)	(0.71)
Share of area planted with legumes (%)	43	33	50	40	42
	(21)	(16)	(25)	(18)	(18)
Farm production diversity score (count)	5.37	3.71	7.99	4.27	5.34
	(4.00)	(2.73)	(4.93)	(3.07)	(3.61)
Gross value of production (000MK/ha)	116	52	129	144	131
	(101)	(48)	(101)	(106)	(105)
Household dietary diversity score	7.83	7.78	7.96	7.53	8.01
(count)	(1.64)	(1.72)	(1.61)	(1.74)	(1.51)
Food consumption score (continuous)	53.00	46.71	79.01	42.47	47.70
	(27.08)	(36.61)	(21.17)	(18.64)	(17.05)
Calorie consumption (Kcal/day/AE)	2973	2981	3387.69	2445	3096
	(1807)	(1828)	(2055)	(1469)	(1756)
Vitamin A consumption (µg	929	693	1109	750	1075
RE/day/AE)	(567)	(563)	(542)	(540)	(517)
Iron consumption (mg/day/AE)	18.04	15.14	21.75	13.31	20.78
	(9.81)	(10.05)	(9.07)	(8.46)	(9.12)
Zinc consumption (mg/day/AE)	16.32	11.48	20.17	10.89	19.25
	(10.14)	(8.34)	(8.85)	(7.06)	(9.64)
Explanatory variable					
HH received subsidized legume seeds	0.07	0.02	0.14	0.07	0.04
(1/0)	(0.25)	(0.14)	(0.34)	(0.26)	(0.19)
Observations (No. of households)	7,034	1,287	1,595	1,813	2,339

Source: Authors' calculations. HH, AE and RE denotes household head, adult equivalent and retinol equivalent, respectively. Mean values are shown with standard deviations in parentheses. Summary statistics for other explanatory variables is shown in Table A6 of the appendix. t-tests were carried out to test for mean differences between beneficiary and non-beneficiary of subsidized legume seeds, and the results are shown in Table A7 in the appendix.

than non-beneficiaries (see, Table A7). This provides early evidence on potential effects of SLSs on farm productivity and dietary quality. Though, these relationships in Tables 2 and A7 do not control for any confounding factors. We address this limitation with panel regression models in the next section.

5.2 Legume seed subsidies and farm productivity

Table 3 presents panel regression results (equation 1) in which area planted with legume crops and VoP are the dependent variables, and household`s access to subsidized legume seeds[2] is the main control variable of interest. Columns (1) and (2) of Table 3 shows that

Table 3: Effects of Subsidizing Legume Seeds on Area Planted to Legume Crops and Value of Production

	Area planted to legume crops (Acres)		Gross value of production (MK/ha)			
			All legume crops		All crops	
	RE	Mundlak	RE	Mundlak	RE	Mundlak
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Subsidized legume seeds						
=1 if HH received subsidized legume seeds	0.197***	0.190***	1.734***	1.701***	0.464***	0.447***
	(0.044)	(0.043)	(0.287)	(0.291)	(0.124)	(0.126)
Household characteristics	Yes	Yes	Yes	Yes	Yes	Yes
Region and year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Mundlak effects	No	Yes	No	Yes	No	Yes
R^2	0.300	0.307	0.113	0.118	0.070	0.073
Observations (No. of households)	6948	6948	6928	6928	6928	6928
Panel B: Lagged subsidized legume seeds						
=1 if HH received subsidized legume seeds	0.215***	0.214***	1.701***	1.664***	0.434***	0.428***
	(0.050)	(0.049)	(0.331)	(0.325)	(0.144)	(0.142)
=1 if HH received subsidized legume seeds (lagged)	0.044	0.037	0.959***	0.876**	0.211	0.188
	(0.043)	(0.043)	(0.357)	(0.361)	(0.220)	(0.215)
Household characteristics	Yes	Yes	Yes	Yes	Yes	Yes
Region and year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Mundlak effects	No	Yes	No	Yes	No	Yes
R^2	0.307	0.312	0.102	0.108	0.034	0.039
Observations (No. of households)	4522	4522	4522	4522	4522	4522

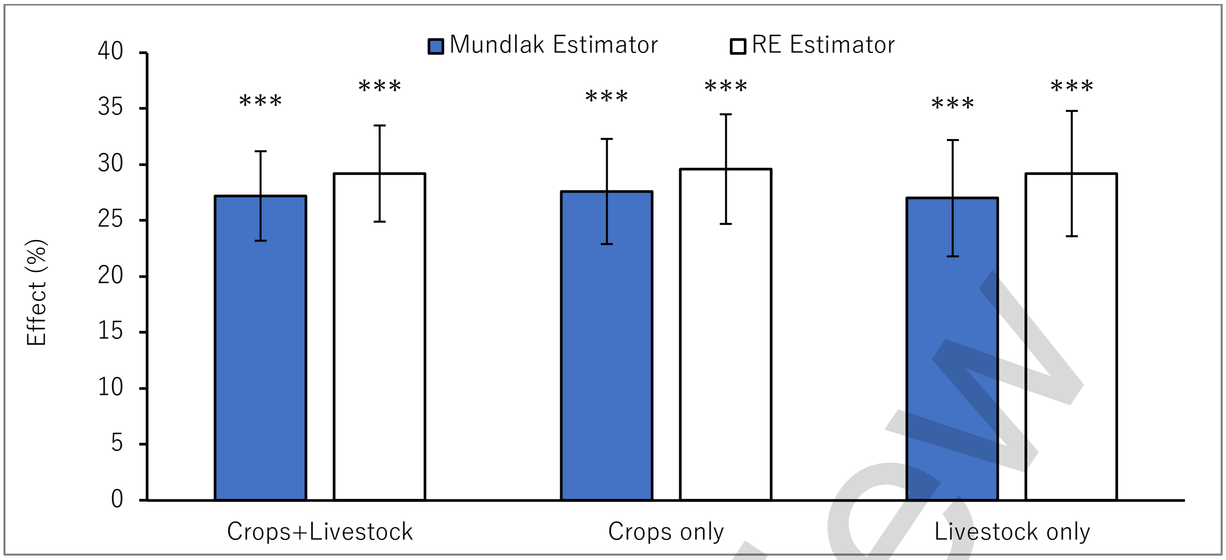
Source: Authors' calculations. To account for zero values for area planted to legume crops and gross value of production, we transformed our dependent variables using inverse hyperbolic sine transformation ($\log(x + (x^2 + 1)^{0.5})$). Coefficient estimates from RE and Mundlak regressions are shown with robust standard errors clustered at enumeration area in parentheses. Full model results are shown in Tables A8 (Panel A) and A9 (Panel B) of the appendix, after controlling for other confounding factors. ** $p < 0{\cdot}05$, *** $p < 0{\cdot}01$.

SLSs increased area planted with legume crops by 19-22%. We also found that past receipt of subsidized legume seeds increases VoP from legume crops in the subsequent season (Panel B of Table 3). As expected, these are substantial immediate outcomes of the subsidy program. Due to shrinking farm sizes in Malawi (Chamberlin & Ricker-Gilbert, 2016), SLSs could be key to improving production through maize-legume intercropping.

As a robustness check, we estimated a fraction probit regression model for columns (1)

[2] Here, we excluded subsidized fertilizer as a control variable in our analysis. Emerging evidence suggests that farmers rarely apply fertilizer to legume crops (Lunduka et al., 2013).

Figure 3: Effects of Subsidizing Legume Seeds on Farm Production Diversity



Source: Authors' calculations. Coefficient estimates—expressed as a percentage: (coefficient–1) * 100—from Poisson estimator through random effects (RE) and Mundlak approaches are shown with standard error bars. Full model results are shown in Table A12 of the appendix. *** p < 0·01. Note: Coefficient estimates from RE and Mundlak regressions are shown with shown with standard error bars. *** p < 0·01. N= 6948.

and (2) of Table 3. We used share of area planted with legume crops as dependent variable and its regression results are shown in Table A10 of the appendix. However, we report its marginal effect results to ease interpretation. We also found that SLSs increased farm land allocated to legume crops by 5-9% (Table A10). Our results suggest that legume seed subsidies are important to increase legume production and farm income, at least among Malawian farmers. Nevertheless, our results are somewhat different from Chibwana et al. (2012) who found that input subsidies reduce farm land allocated to legume crops.

We also estimated equation (1) with gross value of production—as a proxy measure for farm productivity and income—as dependent variable, and its regression results are shown in columns (3) to (6) of Table 3. We found that accessing subsidized legume seeds increased gross value of production by at least 43%. The effect sizes for all legume crops are bigger than all crops. These findings are consistent if we use the control function (CF) approach (see, Table A11 of the appendix).

Figure 3 presents results on the effects of SLSs on farm production diversity. We found that SLSs are positively associated with higher farm production diversity. This implies that accessing subsidized legume seeds leads to an increase in the number of crops grown or livestock species raised by farming households, by at least 27%. However, somehow the magnitude of the effect size is even larger if we use the CF approach (see Table A13 in the appendix). Similarly, Theriault & Smale (2021) found that fertilizer subsidy program crowds in target crops in Mali.

Our findings suggest that SLSs as a policy intervention may significantly contribute to production diversity at farm level. Moreover, our findings are consistent if we use alternative indicators for production diversity: e.g., if we use number of crops grown only, and animal species raised by the household only (Figure 3). Generally, our results are consistent with other studies (e.g., Theriault & Smale 2021; Ricker-Gilbert & Jayne, 2017; Wossen et al., 2017; Awotide et al., 2013) who found that cereal seed and fertilizer subsidies increase productivity for smallholder farmers.

5.3 Legume seed subsidies and nutrition

Table 4 present regression results with HDDS, FCS, calorie and micronutrient consumption levels as dependent variables. Subsidizing legume seeds has a positive and significant effects on all nutrition indicators, except for iron. Overall, we found that SLSs increased dietary diversity, calories, vitamin A and zinc consumption by 4.5-6.6%, 9.5%, 9.9% and 6.6%, respectively. The effect sizes are even larger if we use the CF approach especially for HDDS and calories (Table A15 of the appendix).

On the other hand, we found that subsidized maize seed and subsidized fertilizer are not associated with improved nutrition, in particular micronutrient consumption (see Table A16 of the appendix). With our panel data, emerging evidence suggests that maize seed and fertilizer are not an effective strategy to improve nutritional outcomes such as vitamin A and zinc consumption. This implies that addressing malnutrition, especially micronutrient deficiencies may require African governments and development partners to invest more in legume seed subsidies than maize seed subsidies.

We further examined the bivariate relationship between SLSs and calories/micronutrient consumption using predicted values from Table 4 to better understand continuous treatment effects. Results are shown in Figure A2 of the appendix, where kernel density of calories, vitamin A, iron and zinc by smallholder farming households who either had accessed

Table 4: Effects of Subsidizing Legume Seeds on Dietary Diversity, Calorie and Micronutrient Consumption

	HDDS	FCS	Calories	Vitamin A	Iron	Zinc
	(1)	(2)	(3)	(4)	(5)	(6)
=1 if HH received subsidized legume seeds	1.045***	0.066**	0.095***	0.099*	0.021	0.066*
	(0.010)	(0.029)	(0.030)	(0.056)	(0.040)	(0.039)
Household characteristics	Yes	Yes	Yes	Yes	Yes	Yes
Region fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Mundlak effects	Yes	Yes	Yes	Yes	Yes	Yes
Pseudo R^2/R^2	0.017	0.250	0.133	0.153	0.221	0.282
Observations	6928	6928	6825	6825	6825	6825

Source: Authors' calculations. HDDS and FCS, denotes household dietary diversity score and food consumption score, respectively. Calories and micronutrient—e.g., vitamin A, iron and zinc—consumption are expressed as logarithm. Coefficient estimates from Mundlak regressions are shown with robust standard errors clustered at enumeration area in parentheses. Full model results are shown in Table A14 of the appendix. The null hypothesis on access to subsidized legume seeds being exogenous could not be rejected (except for HDDS and Calories regression models, which are presented in Table A15 of the appendix), so that panel linear regression model estimates are shown. * p < 0·10, ** p < 0·05, *** p < 0·01.

subsidized legume seeds or not are plotted. For all dietary quality indicators, SLSs continuously increased consumption of vitamin A, iron, and zinc even after controlling for confounding factors.

To obtain more precise measures of nutrition, we also used two child anthropometric outcomes: WAZ and HAZ, which are calculated using WHO's child growth standard (O'Donnell et al., 2008). We used anthropometric data for 7971 children aged between 0 and 18 years. We focused on WAZ and HAZ as indicators for short and long-run nutrition outcomes, respectively. Regression results for child WAZ and HAZ models are presented in Tables A17 and A18 of the appendix, respectively. Consistent with findings from food consumption data (i.e., Table 4 results), we further found that SLSs is positively correlated with WAZ but not HAZ. These results further suggest that legume seed subsidies could improve short-term nutritional outcomes, especially among children. These are welcome findings as undernutrition and micronutrient deficiencies are common in SSA (Zaharia et al., 2021; Kinyoki et al., 2020; Unicef et al., 2019).

Our findings suggest that SLSs as a policy intervention could be key to improving farm productivity, household diets and nutrition for several reasons. First, unlike cereal seeds like maize, smallholder farmers often plant local or traditional recycled legume seed varieties. If targeted properly, SLSs may help to create demand for improved legume seeds as well as increasing its adoption in the long-run. This may increase productivity of legume crops. Second, legumes are often a cheap source for protein to most smallholder farmers than ASFs. Third, grains of legumes or their leaves are more nutritious than those of most cereals. Finally, most legume crops attract relatively good farm-gate prices, and they generally require low amount of fertilizers than cereals such as maize, and this may increase farm income from legume crops.

5.4 Additional analyses

To further check the consistency of our main results, we estimated two sets of alternative models. First, we used quantities (as opposed to a binary variable) of subsidized legume seeds received by the household as a key explanatory variable of interest (intensive margins). Full model results are shown in Tables A19 and A20 of the appendix. The results are still consistent with our main results, suggesting that SLSs is positively correlated with area planted with legume crops, gross value of production, farm production diversity, consumption (dietary) diversity, calories, and vitamin A consumption.

Due to land scarcity, there are some very small fields in Malawi, which could magnify measurement errors when calculating gross value of production. Moreover, most productivity indicators may suffer from measurement errors especially if farmers use recall data to capture farm size (Wollburg et al., 2021). Hence, we also examined the robustness of our findings by splitting the sample into three farm size categories: less than 0.2 ha, 0.2-2.5 ha, and greater than 2.5 ha. The results are reported in Table A21 of the appendix. Consistent with our main results in Table 3, Table A21 results are robust to the inclusion of very small plots.

5.5 Limitations of the study

There are several limitations in this study. First, our results are specific for Malawi, but their policy relevance could be generalized beyond most case studies with cross-sectional datasets, especially countries with similar conditions. Second, with observational data, it is extremely difficult to have perfect causal results on linkage between legume seed subsidies and farm productivity or nutrition. Hence, our results should be cautiously interpreted as associations.

Lastly, with limited sample size—out of 7034 households in our sample, it was only 7% who had acquired subsidized legume seeds—for households who had accessed subsidized legume seeds, it could be interesting for future research to validate our results, especially on child nutrition with a larger sample size and in a different context. Moreover, to the best of our knowledge, we have not seen any study that has analyzed the linkages between legume seed subsidies and productivity or nutrition in an experimental setting.

6. Conclusion and policy implications

Using a decade-long nationally representative panel data from smallholder farmers in a developing country (Malawi), we have shown that subsidizing legume seeds (SLSs) is associated with positive effects on area planted with legume crops, farm productivity, dietary quality, and child nutrition. With widespread hunger, undernutrition, and micronutrient malnutrition as well as poverty in many developing countries, particularly SSA (Zaharia et al., 2021; Kinyoki et al., 2020; Unicef et al., 2019). The need to reverse this situation through productivity-enhancing agricultural technologies such as improved seeds and inorganic fertilizers and pro-nutrition technologies is paramount.

Consequently, many countries in SSA are implementing agricultural ISPs, especially for improved cereal seeds and inorganic fertilizers, and legume seeds to a lesser extent. However, it remains largely unknown as to whether SLSs can improve nutrition through productivity and income pathways or not. Most existing studies (e.g., Theriault & Smale, 2021; Abman & Carney, 2020; Ricker-Gilbert & Jayne, 2017) have largely analyzed the role of input—cereal seeds and inorganic fertilizers—subsidies on crop species diversity, crop yields, income, and poverty, but not on dietary quality and child nutrition. Moreover, other similar studies (Theriault & Smale, 2021; Smale et al., 2020; Chibwana et al., 2012) on ISPs used datasets that were cross-sectional where controlling for possible unobserved confounding factors is difficult, and these studies were also not nationally representative.

We have addressed this important gap by analyzing the effects of SLSs on farm productivity and nutrition using nationally representative panel data from Malawian smallholder farmers. We have used panel regression models with instrumental variable approach to address potential endogeneity issues. We found evidence that SLSs increased area planted with legume crops, gross value of production, both production and consumption (dietary) diversity, calories, vitamin A and zinc consumption. We further found that SLSs was positively correlated with child WAZ but not HAZ. The positive effects on dietary diversity, micronutrient consumption and child nutrition, in particular child WAZ, are welcome findings as malnutrition remains a widespread public health problem in many developing countries.

Although most countries implementing ISPs in SSA pay more attention to cereal seeds only, our findings suggest that SLSs as a policy intervention could have substantial benefits on farm productivity and nutrition. If targeted appropriately, SLSs could improve household nutrition through both productivity and income pathways. Inclusion of legume seeds in an ISP, not only could increase adoption of improved legume seeds, but also increase crop yields and income from legume crops. On the other hand, allocating more resources to maize seed and fertilizer subsidies, might not be an effective strategy in addressing malnutrition, particularly micronutrient deficiencies in SSA. Nevertheless, diversifying (or adopting a flexible) input subsidy portfolios—e.g., pesticides/herbicides, organic inputs, and livestock drugs, beyond pro-nutrition seed varieties and inorganic fertilizers is required for sustaining the benefits of ISPs.

Acknowledgements

This research was produced with the help of a grant from the Foundation for a Smoke-Free World (FSFW) Agricultural Transformation Initiative (ATI) through the Michigan State University (MSU) Food Security Group. We are grateful for this generous support. Any views expressed herein are those of the authors and do not necessarily represent the views of the donors.

References

Abman, R. and Carney, C. 2020. Agricultural productivity and deforestation: Evidence from input subsidies and ethnic favouritism in Malawi. *Journal of Environmental Economics and Management,* 103, 102342.

Ariga, J., Mabaya, E., Waithaka, M. and Wanzala-Mlobela, M. 2019. Can improved agricultural technologies spur a green revolution in Africa? A multicountry analysis of seed and fertilizer delivery systems. *Agricultural Economics,* 50, 63–74.

Awotide, B.A., Karimoy, A., Diagne, A. and Nakelse, T. 2013. The impact of seed vouchers on poverty reduction among smallholder rice farmers in Nigeria. *Agricultural Economics, 44,* 647–658.

Benson, T. 2021. *Disentangling food security from subsistence agriculture in Malawi*. Washington, DC: International Food Policy Research Institute.

Chamberlin, J. and Ricker-Gilbert, J., 2016. Participation in rural land rental markets in Sub - Saharan Africa: Who benefits and by how much? Evidence from Malawi and Zambia. *American Journal of Agricultural Economics,* 98 (5), 1507–1528

Chibwana, C., Fisher, M. and Shively, G. 2012. Cropland allocation effects of agricultural input subsidies in Malawi. *World Development,* 40 (1), 124–133.

Chilundu, S. 2020. "Affordable Inputs Programme to use biometric system-Lowe". The Nation, August 2, 2020 edition.

Debela, B.L., Gehrke, E. and Qaim, M. 2021. Links between maternal employment and child nutrition in rural Tanzania. *American Journal of Agricultural Economics,* 103 (3), 812–830.

Development Initiatives. 2018. "Global Nutrition Report 2018". Bristol: Development Initiatives.

Development Initiatives. 2020. "Global Nutrition Report 2020 Country Nutrition Profiles". Bristol: Development Initiatives. Accessed 3/26/2021 at: https://globalnutritionreport.org/resources/nutrition-profiles/africa/eastern-africa/malawi/.

Ecker, O. 2018. Agricultural transformation and food and nutrition security in Ghana: Does farm production diversity (still) matter for household dietary diversity? *Food Policy*, 79*,* 271–282.

FAO, IFAD, UNICEF, WFP and WHO. 2019. *The state of food security and nutrition in the World 2019: Safeguarding against economic slowdowns and downturns*. Rome, FAO.

FAOSTAT. 2021. Accessed August 2021, at http://www.fao.org/faostat/en/.

Harou, A.P. 2018. Unraveling the effect of targeted input subsidies on dietary diversity in household consumption and child nutrition: The case of Malawi. *World Development,* 106*,* 124–135.

Hirvonen, K., Bai, Y., Headey, D. and Masters, W.A 2020. Affordability of the EAT–Lancet reference diet: A global analysis. *Lancet Global Health,* 8*,* e59–66.

Hoddinott, J., Sandström, S. and Upton, J. 2018. The impact of cash and food transfers: Evidence from a randomized intervention in Niger. *American Journal of Agricultural Economics,* 100 (4), 1032–1049.

Holden, S. and Lunduka, R. 2012. Do fertilizer subsidies crowd out organic manures? The case of Malawi. *Agricultural Economics,* 43, 303–314.

Khonje, M.G., Ecker, O. and Qaim, M. 2020. Effects of modern food retailers on adult and child diets and nutrition. *Nutrients,* 12, 1714.

Kinyoki, D.K., et al. 2020. Mapping child growth failure across low-and middle-income countries. *Nature,* 577, 231–234.

Lunduka, R., Ricker-Gilbert, J. and Fisher, M. 2013. What are the farm-level impacts of Malawi's farm input subsidy program? A critical review. *Agricultural Economics,* 44, 563–579.

Mason, N.M. and Ricker-Gilbert, J. 2013. Disrupting demand for commercial seed: Input subsidies in Malawi and Zambia. *World Development,* 45, 75–91.

Mason, N.M. and Smale, M. 2013. Impacts of subsidized hybrid seed on indicators of economic well-being among smallholder maize growers in Zambia. *Agricultural Economics,* 44, 659–670.

Mundlak, Y. 1978. On the pooling of time series and cross-section data. *Econometrica,* 46, 69–85.

Muthini, D., Nzuma, J. and Qaim, M. 2020. Subsistence production, markets, and dietary diversity in the Kenyan small farm sector. *Food Policy,* 97, 101956.

Muyanga, M. and Jayne, T.S. 2019. Revisiting the farm size-productivity relationship based on a relatively wide range of farm sizes: Evidence from Kenya. *American Journal of Agricultural Economics,* 101, 1140–1163.

O'Donnell, O., van Doorslaer, E., Wagstaff, A. and Lidelow, M. 2008. *Analyzing health equity using household survey data: A guide to techniques and their implementation*. Washington DC: World Bank.

Ogutu, S. O., Gödecke, T. and Qaim, M. 2019. Agricultural commercialisation and nutrition in smallholder farm households. *Journal of Agricultural Economics,* 71 (2), 534–555.

Ogutu, S.O., Fongar, A., Gödecke, T., Jäckering, L., Mwololo, H., Njuguna, M., Wollni, M. and Qaim, M. 2020. How to make farming and agricultural extension more nutrition-sensitive: Evidence from a randomised controlled trial in Kenya. *European Review of Agricultural Economics,* 47 (1), 95–118.

Ricker-Gilbert, J. and Chamberlin, J. 2018. Transaction costs, land rental markets, and their impact on youth access to agriculture in Tanzania. *Land Economics,* 94 (4), 541–555.

Ricker-Gilbert, J. and Jayne, T.S. 2017. Estimating the enduring effects of fertiliser subsidies on commercial fertiliser demand and maize production: Panel data evidence from Malawi. *Journal of Agricultural Economics,* 68 (1), 70–97.

Rubyogo, J.C., et al. 2019. Market-led options to scale up legume seeds in developing countries: Experiences from the Tropical Legumes Project. *Plant Breeding,* 138 (4), 474–486.

Sibhatu, K.T., Krishna, V.V. and Qaim, M. 2015. Production diversity and dietary diversity in smallholder farm households. *Proceedings of the National Academy of Sciences of the United States of America,* 112, 10657–10662.

Smale, M., Thériault, V. and Mason, N.M. 2020. Does subsidizing fertilizer contribute to the diet quality of farm women? Evidence from rural Mali. *Food Security,* 12, 1407–1424.

Snapp, S., Jayne, T. S., Mhango, W., Benson, T. and Ricker-Gilbert, J. 2014. *Maize-nitrogen response in Malawi's smallholder production systems*. Malawi strategy support program working paper.

Theriault, V. and Smale, N.M. 2021. The unintended consequences of the fertilizer subsidy program on crop species diversity in Mali. *Food Policy,* 102, 102121.

Tittonell, P. and Giller, K.E. 2013. When yield gaps are poverty traps: The paradigm of ecological intensification in African smallholder agriculture. *Field Crops Research,* 143, 76–90.

UNICEF, WHO and World Bank. 2019. *Levels and trends in child malnutrition: Key findings of the 2019 edition of the joint child malnutrition estimates*. Geneva: WHO.

Villa, K.M., Barrett, C.B. and Just, D.R. 2011. Whose fast and whose feast? Intrahousehold asymmetries in dietary diversity response among East African Pastoralists. *American Journal of Agricultural Economics,* 93 (4), 062–1081.

Walls, H.L., Johnston, D., Tak, M., Dixon, J., Hanefeld, J., Hull, E. and Smith, R.D. 2018. The impact of agricultural input subsidies on food and nutrition security: A systematic review. *Food Security,* 10, 1425–1436.

Wollburg, P., Tiberti, M. and Zezza, A. 2021. Recall length and measurement error in agricultural surveys*. Food Policy,* 100, 102003.

Wooldridge, J.M. 2019. Correlated random effects models with unbalanced panels. *Journal of Econometrics,* 211, 137–150.

World Food Programme (WFP). 2008. *Calculation and use of the food consumption score in food security analysis*. Vulnerability analysis and mapping branch, WFP.

Wossen, T., Abdoulaye, T., Alene, A., Feleke, S., Ricker-Gilbert, J., Manyong, V. and Awotide, B.A. 2017. Productivity and welfare effects of Nigeria's e-voucher-based input subsidy program. *World Development,* 97, 251–265.

Zaharia, S., et al. 2021. Sustained intake of animal-sourced foods is associated with less stunting in young children. *Nature Food,* 2, 246–254.

Zezza, A., Carletto, C., Fiedler, J. L., Gennari, P. and Jolliffe, D. 2017. Food counts: Measuring food consumption and expenditures in household consumption and expenditure surveys. Introduction to the special issue. *Food Policy,* 72, 1–6.

Appendix – Supplementary Tables and Figures

Contents

Table A1. Subsidized input supply by survey year in Malawi

Subsidized inputs supplied	2010/11	2013/14	2016/17	2019/20	Total
	(1)	(2)	(3)	(4)	(5)
Fertilizer (MT)	**159,952**	**149,971**	**89,511**	**89,880**	**489,314**
NPK	79,945	74,991	44,776	44,944	244,656
Urea	80,007	74,979	44,735	44,936	244,657
Maize (MT)	**10,650**	**8,268**	**4,628**	**4,472**	**28,018**
Hybrid	8,521	6,087	3,129	3,345	21,082
OPV	2,129	2,181	1,499	1,127	6,936
Legume (MT)	**2,727**	**3,042**	**1,664**	**1,283**	**8,715**
Groundnuts	2,029	2,151	776	138	5095
Soybeans	375	384	327	785	1871
Beans	316	475	438	324	1554
Pigeon peas	4	17	98	0	119
Cowpeas	2	14	25	35	76

Note: MT, metric tonnes. The data were drawn from Government of Malawi`s final report on the implementation of Agricultural Inputs Subsidy Program for 2010/11, 2013/14, 2016/17 and 2019/20 growing seasons.

Table A2. Percentage of sampled farming households growing major crops by survey year

	All panel years	2010	2013	2016	2019
	(1)	(2)	(3)	(4)	(5)
Maize	95.38	97.99	95.83	94.63	94.25
Groundnuts	30.84	32.56	36.81	21.46	33.42
Pigeon peas	24.56	20.74	27.25	20.19	28.32
Nkhwani (Common bean leaves)	24.47	6.99	18.78	15.71	44.53
Common beans	11.21	6.19	10.50	9.35	15.86
Sorghum	11.06	9.89	10.77	11.67	11.42
Tobacco	10.67	15.43	11.04	10.18	8.24
Soy beans	10.61	6.43	10.43	10.67	12.94
Sweet potatoes	6.51	4.74	4.17	4.98	10.20
Peas	3.53	0.96	6.80	2.49	3.62
Rice	3.38	3.62	3.63	2.21	4.01
Cotton	2.09	0.80	4.44	2.21	1.18
Pearl millet	1.57	1.13	1.68	1.27	1.96
Tomato	1.32	1.77	1.35	0.55	1.66
Sunflower	1.23	1.29	1.35	1.22	1.13
Okra	1.11	0.72	1.28	0.77	1.48
Finger millet	1.05	0.80	0.81	1.33	1.13
Ground beans	0.82	0.88	1.21	0.39	0.87
Irish potato	0.78	0.80	0.54	0.50	1.13
Sugarcane	0.48	0.16	0.07	0.39	1.00
Tanaposi	0.19	0.08	0.20	0.17	0.26
Paprika	0.18	0.08	0.27		0.17
Wheat	0.10	0.16		0.06	
Onion	0.06	0.08	0.07		0.04
Others	4.73	2.33	4.85	3.26	7.10
Observations (Number of households)	6,833	1,244	1,486	1,808	2,295

Table A3. Food groups/items and weights used for construction of dietary diversity indicators

Panel A: Dietary diversity score

Group	Food Items (examples only)	Food Group	Weight
1	Maize, rice, sorghum, millet, bread, cassava, potatoes, and plantains	Cereals	1
2	Cassava, potatoes, and plantains	Tubers and roots	1
3	Beans, cowpea, groundnuts, pigeonpea, soybean, and velvet beans	Legumes and nuts	1
4	Beef, chicken, ducks, goat meat, sheep meat, and pork	Meat	1
5	Fish (fresh/frozen/dried), and tinned fish	Fish and fish products	1
6	Eggs	Eggs	1
7	Vegetables, and leaves (i.e., pulses) Cassava leaves, and sweet potato leaves	Vegetables	1
8	Fruits	Fruits	1
9	Milk, yoghurt, and other dairy products	Milk and milk products	1
10	Sweets, sugar, tea, coffee, soft drinks	Sweets, sugars and beverages	1
11	Cooking oil	Oils and fats	1
12	Salt, spices, sauce	Spices and condiments	1

Panel B: Food consumption score (FCS)

Group	Food Items (examples only)	Food Group	Weight
1	Maize, maize porridge, rice, sorghum, millet, bread, cassava, potatoes, and sweet potatoes	Staples	2
2	Beans, peas, and groundnuts	Pulses	3
3	Vegetables, and leaves (i.e., pulses)	Vegetables	1
4	Fruits	Fruits	1
5	Beef, goat, poultry, pork, eggs, and fish	Meat and fish	4
6	Milk, yoghurt, and other dairy products	Dairy	4
7	Sugar, sugar products, and honey	Sugar	0.5
8	Oils, fats, and butter	Oil	0.5

Note: The components and scoring standards for FCS are adopted from WFP (2008).

Table A4. District-level population shares by survey year

	2010	2013	2016	2019
	(1)	(2)	(3)	(4)
Chitipa	1.36	1.34	1.32	1.30
Karonga	2.07	2.07	2.07	2.06
Nkhatabay	1.65	1.65	1.65	1.65
Rumphi	1.31	1.29	1.27	1.25
Mzimba	5.54	5.51	5.48	5.46
Likoma	0.07	0.07	0.06	0.06
Mzuzu City	1.12	1.27	1.42	1.56
Kasungu	4.88	4.99	5.10	5.21
Nkhotakota	2.33	2.33	2.33	2.33
Ntchisi	1.73	1.75	1.76	1.76
Dowa	4.40	4.58	4.74	4.87
Salima	2.59	2.58	2.57	2.55
Lilongwe	9.28	9.06	8.86	8.66
Mchinji	3.54	3.59	3.63	3.67
Dedza	4.70	4.59	4.47	4.36
Ntcheu	3.58	3.54	3.49	3.44
Lilongwe City	5.51	6.02	6.52	7.00
Mangochi	6.13	6.19	6.26	6.34
Machinga	3.75	3.73	3.73	3.73
Zomba	4.32	4.16	4.00	3.85
Chiradzulu	2.13	2.02	1.92	1.82
Blantyre	2.56	2.49	2.42	2.36
Mwanza	0.69	0.66	0.63	0.60
Thyolo	4.26	4.07	3.89	3.74
Balaka	2.43	2.43	2.43	2.44
Chikwawa	3.31	3.29	3.27	3.25
Neno	0.85	0.89	0.94	0.98
Phalombe	2.37	2.32	2.28	2.24
Blantyre City	5.17	5.33	5.47	5.58
Mulanje	3.85	3.64	3.44	3.26
Nsanje	1.79	1.75	1.71	1.68
Zomba City	0.72	0.80	0.87	0.94
Total population	13,947,592	15,316,860	16,832,910	18,508,613

Table A5. Testing validity of instrument variable

Dependent variable	=1 if HH received subsidized legume seeds	Gross value of production (All legumes)	FCS	Calories	Vit. A	Iron	Zinc
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Population share	–0.009***	–0.022	0.153	–7.655	3.222	–0.040	0.017
	(0.010)	(0.032)	(0.132)	(5.469)	(2.913)	(0.050)	(0.051)
Hh characteristics	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R-squared		0.061	0.198	0.073	0.123	0.153	0.155
Observations (No. of Hhs)	6928	6928	6928	5100	6825	6825	6825

Note: Hh, households. Vit.A, Vitamin A. Coefficient estimates from Generalized Linear Model and OLS regressions are shown in parentheses. We excluded households with calorie consumption of more than 4000 as outliers. Other covariates were included for estimation but are not shown here for brevity. *** $p < 0{\cdot}01$.

Table A6. Descriptive results for control variables by survey year

	All panel years	2010	2013	2016	2019
	(1)	(2)	(3)	(4)	(5)
Explanatory variables					
HH received subsidized legume seeds (1/0)	0.07	0.02	0.14	0.07	0.04
	(0.25)	(0.14)	(0.34)	(0.26)	(0.19)
Male household head (1/0)	0.74	0.77	0.76	0.74	0.71
	(0.44)	(0.42)	(0.43)	(0.44)	(0.45)
Age of household head (years)	43.46	42.61	43.48	43.80	43.65
	(15.92)	(16.09)	(15.83)	(15.86)	(15.92)
Assets value (MK/year)	129990	29709	90258	122288	218233
	(981186)	(158221)	(633224)	(681057)	(1494965)
Total cultivated land (ha)	1.80	2.03	1.80	1.77	1.69
	(1.85)	(2.72)	(1.62)	(1.70)	(1.48)
Access to credit (1/0)	0.22	0.12	0.20	0.24	0.27
	(0.41)	(0.33)	(0.40)	(0.43)	(0.44)
Access to extension (1/0)	0.46	0.00	0.01	0.76	0.80
	(0.50)	(0.03)	(0.10)	(0.43)	(0.40)
Observations (No. of households)	7034	1287	1595	1813	2339

Note: HH, household head. Mean values are shown with standard deviations in parentheses. t-tests were carried out to test for mean differences between non-beneficiary and beneficiary of subsidized legume seeds, and the results are shown in Table A7.

Table A7. Descriptive statistics by recipient of subsidized legume seeds

	Did the HH receive subsidized legume seeds?		Mean difference
	Yes	No	
	(1)	(2)	(3)= (1)–(2)
Dependent variables			
Area planted with legume crops (acres)	0.82	0.55	0.27***
	(0.75)	(0.69)	(0.02)
Share of area planted with legume crops (%)	31	24	6***
	(26)	(26)	(0.64)
Farm production diversity score (count)	8.66	5.14	3.52***
	(4.97)	(3.82)	(0.19)
Gross value of production (000MK/ha)	138	115	23***
	(96)	(101)	(2)
Household dietary diversity score (count)	8.17	7.81	0.36***
	(1.29)	(1.66)	(0.08)
Food consumption score (continuous)	62.79	52.32	10.47***
	(25.14)	(27.08)	(1.30)
Calories consumption (Kcal/day/AE)	3257.12	2953.31	303.81***
	(1915.66)	(1797.37)	(87.92)
Vitamin A consumption (μg RE/day/AE)	1028.73	922.04	106.69***
	(567.16)	(566.30)	(27.58)
Iron consumption (mg/day/AE)	19.74	17.92	1.82***
	(10.26)	(9.76)	(0.48)
Zinc consumption (mg/day/AE)	18.50	16.16	2.33***
	(10.67)	(10.09)	(0.49)
Child nutrition			
Weight-for-age Z-scores (WAZ)	–0.32	–0.47	–0.15**
	(1.16)	(1.18)	(0.06)
Height-for-age Z-scores (HAZ)	–0.90	–1.01	–0.11
	(1.52)	(1.50)	(0.07)
Explanatory variables			
Male HH (1/0)	0.75	0.74	0.01
	(0.43)	(0.44)	(0.02)
Age of HH (years)	45.48	43.31	2.17***
	(16.20)	(15.89)	(0.77)
Assets value (MWK/year)	73191	133974	–60783
	(501220)	(1006196)	(47272)
Total cultivated land (ha)	2.12	1.78	0.35***
	(1.60)	(1.86)	(0.09)
Access to credit (1/0)	0.23	0.22	0.01
	(0.42)	(0.41)	(0.02)
Access to extension (1/0)	0.39	0.47	–0.07***
	(0.49)	(0.50)	(0.02)
Observations (Number of households)	461	6573	7034

Note: HH, AE and RE denotes household head, adult equivalent and retinol equivalent, respectively. Mean values are shown with standard deviations or (standard errors for column 3) in parentheses. t-tests were carried out to test for mean differences between non-beneficiary and beneficiary of subsidized legume seeds. ** $p < 0.05$, *** $p < 0.01$.

Table A8. Effects of subsidizing legume seeds on area planted to legume crops and value of production: full model results in Table 3, Panel A

	Area planted to legume crops (Acres)		Gross value of production (MK/ha)			
			All legumes		All crops	
	RE	Mundlak	RE	Mundlak	RE	Mundlak
=1 if HH received subsidized legume seeds	0.197***	0.190***	1.734***	1.701***	0.464***	0.447***
	(0.044)	(0.043)	(0.287)	(0.291)	(0.124)	(0.126)
Male household head	-0.067***	-0.074***	–0.217	–0.258	0.070	0.089
	(0.022)	(0.022)	(0.228)	(0.223)	(0.127)	(0.126)
Age of the household head (log)	-0.064	-0.075	–0.017	–0.574	0.089	–1.015**
	(0.073)	(0.094)	(0.625)	(0.795)	(0.352)	(0.493)
Age of household head squared	0.000**	0.000*	0.000	0.000	0.000	–0.000**
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Literacy (dummy)	0.03	0.036*	0.084	0.137	–0.033	–0.001
	(0.023)	(0.022)	(0.206)	(0.202)	(0.093)	(0.091)
Asset value (log)	0.002	0.002	0.023	0.029	0.031**	0.032**
	(0.003)	(0.003)	(0.027)	(0.027)	(0.014)	(0.014)
Credit (dummy)	0.007	0.007	0.341**	0.348**	0.131	0.132
	(0.020)	(0.020)	(0.161)	(0.162)	(0.090)	(0.091)
Extension (dummy)	-0.017	-0.023	0.062	0.038	0.541***	0.535***
	(0.024)	(0.024)	(0.275)	(0.274)	(0.184)	(0.187)
Total farm size (ha)	0.208***	0.169***	0.737***	0.532***	0.084	0.006
	(0.013)	(0.012)	(0.086)	(0.080)	(0.054)	(0.045)
Farm size squared	-0.004***	-0.004***	–0.018***	–0.015***	–0.002	–0.001
	(0.000)	(0.000)	(0.003)	(0.002)	(0.001)	(0.001)
Rainfall (log)	0.002	0.027**	0.018	–0.002	–0.060***	0.036
	(0.004)	(0.011)	(0.050)	(0.121)	(0.021)	(0.085)
North (dummy)	-0.171**	-0.102	–1.982***	–1.948***	–1.305***	–1.107**
	(0.087)	(0.085)	(0.738)	(0.724)	(0.453)	(0.456)
South (dummy)	-0.272***	-0.262***	0.377	0.44	–0.276	–0.233
	(0.042)	(0.043)	(0.466)	(0.465)	(0.218)	(0.232)
2013 (dummy)	0.708***	0.705***	1.468***	1.506***	0.807***	0.886***
	(0.043)	(0.042)	(0.274)	(0.277)	(0.111)	(0.119)
2016 (dummy)	0.113***	0.112***	0.253	0.334	0.655**	0.876***
	(0.031)	(0.032)	(0.343)	(0.357)	(0.266)	(0.268)
2019 (dummy)	0.402***	0.393***	3.463***	3.575***	1.410***	1.714***
	(0.037)	(0.041)	(0.392)	(0.437)	(0.240)	(0.254)
Mundlak (time averages) test (□2)		31.440***		20.210***		27.800***
Age of the household head (mean)		0.000		0.024		0.053***
		(0.003)		(0.025)		(0.014)
Asset value (mean)		0.000		–0.000**		0.000
		(0.000)		(0.000)		(0.000)
Total farm size (mean)		0.062***		0.404***		0.153***
		(0.013)		(0.108)		(0.058)
Rainfall (mean)		-0.000**		0.000		–0.001
		(0.000)		(0.001)		(0.001)
Constant	0.337	0.346	3.528	4.780*	10.397***	13.019***
	(0.284)	(0.316)	(2.515)	(2.815)	(1.353)	(1.643)
Wald □2	779***	864***	557***	632***	320***	399***
Observations (No. of households)	6948	6948	6928	6928	6928	6928

Note: We transformed our dependent variables using inverse hyperbola sine (HIS) transformation ($\log(x + (x^2 + 1)^{0.5})$). Coefficient estimates from RE and Mundlak regressions are shown with robust standard errors clustered at enumeration area in parentheses. * p < 0·10, ** p < 0·05, *** p < 0·01.

Table A9. Effects of subsidizing legume seeds on area planted to legume crops and value of production: full model results in Table 3, Panel B

	Area planted to legume crops (Acres)		Gross value of production (MK/ha)			
			All legumes		All crops	
	RE	Mundlak	RE	Mundlak	RE	Mundlak
=1 if HH received subsidized legume seeds	0.215***	0.214***	1.701***	1.664***	0.434***	0.428***
	(0.050)	(0.049)	(0.331)	(0.325)	(0.144)	(0.142)
=1 if HH received subsidized legume seeds (lagged)	0.044	0.037	0.959***	0.876**	0.211	0.188
	(0.043)	(0.043)	(0.357)	(0.361)	(0.220)	(0.215)
Male household head	-0.075***	-0.073***	-0.144	-0.164	0.033	0.075
	(0.026)	(0.025)	(0.260)	(0.250)	(0.138)	(0.132)
Age of the household head (log)	0.038	-0.142	-0.157	-0.935	0.276	-0.966
	(0.090)	(0.107)	(0.814)	(1.028)	(0.453)	(0.607)
Age of household head squared	0.000	0.000	0.000	0.000	0.000	-0.000**
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Literacy (dummy)	0.042	0.049*	0.195	0.248	0.124	0.148
	(0.026)	(0.025)	(0.232)	(0.230)	(0.121)	(0.122)
Asset value (log)	0.002	0.002	0.048*	0.057**	0.041**	0.042**
	(0.003)	(0.003)	(0.028)	(0.028)	(0.017)	(0.017)
Credit (dummy)	0.024	0.027	0.350*	0.381*	0.008	0.03
	(0.025)	(0.025)	(0.202)	(0.205)	(0.111)	(0.114)
Extension (dummy)	-0.021	-0.027	0.138	0.114	0.492**	0.496**
	(0.027)	(0.027)	(0.327)	(0.329)	(0.219)	(0.225)
Total farm size (ha)	0.242***	0.208***	0.715***	0.531***	-0.024	-0.059
	(0.021)	(0.022)	(0.136)	(0.129)	(0.078)	(0.067)
Farm size squared	-0.006***	-0.006***	-0.023**	-0.019**	0.002	0.002
	(0.002)	(0.002)	(0.010)	(0.009)	(0.003)	(0.002)
Rainfall (log)	0.023	0.046	0.534	0.131	0.25	0.1
	(0.030)	(0.039)	(0.507)	(0.503)	(0.246)	(0.278)
North (dummy)	-0.231**	-0.206*	-2.495***	-2.770***	-1.218***	-1.369***
	(0.108)	(0.108)	(0.852)	(0.898)	(0.445)	(0.470)
South (dummy)	-0.313***	-0.307***	0.248	0.291	-0.277	-0.257
	(0.053)	(0.053)	(0.499)	(0.492)	(0.234)	(0.239)
2013 (dummy)	0.351***	0.306***	-1.804***	-1.833***	-0.451	-0.655*
	(0.049)	(0.051)	(0.469)	(0.491)	(0.284)	(0.344)
2016 (dummy)	-0.284***	-0.288***	-3.026***	-3.275***	-0.524***	-0.718***
	(0.033)	(0.036)	(0.390)	(0.424)	(0.180)	(0.218)
Mundlak (time averages) test (□2)		19.030***		21.980***		15***
Age of the household head (mean)		0.009**		0.041		0.063***
		(0.004)		(0.036)		(0.019)
Asset value (mean)		-0.000**		-0.000***		0.000
		(0.000)		(0.000)		(0.000)
Total farm size (mean)		0.052***		0.348***		0.069
		(0.015)		(0.118)		(0.070)
Rainfall (mean)		0.000		0.001		0.001
		(0.000)		(0.001)		(0.001)
Constant	0.149	0.518	3.425	7.027	8.801***	12.780***
	(0.446)	(0.428)	(5.033)	(5.037)	(2.351)	(2.390)
Wald □2	668	723	340	404	122	127
Observations (No. of households)	4522	4522	4522	4522	4522	4522

Note: We transformed our dependent variables using inverse hyperbola sine (HIS) transformation ($\log(x + (x^2 + 1)^{0.5})$). Coefficient estimates from RE and Mundlak regressions are shown with robust standard errors clustered at enumeration area in parentheses. * p < 0·10, ** p < 0·05, *** p < 0·01.

Table A10. Effects of subsidizing legume seeds on share of area planted with legume crops (fractional probit regression model)

	Share of area planted with legume crops			
	(1)	(2)	(3)	(4)
=1 if HH received subsidized legume seeds	0.282***	0.496***	0.366***	0.359***
	(0.060)	(0.059)	(0.059)	(0.059)
=1 if HH received subsidized legume seeds (marginal effects)	*0.053****	*0.090****	*0.064****	*0.063****
	(0.011)	*(0.011)*	*(0.010)*	*(0.010)*
Male household head	–0.123***	–0.232***	–0.175***	–0.173***
	(0.040)	(0.040)	(0.039)	(0.040)
Age of the household head (log)	–0.216	–0.316**	–0.349***	–0.503**
	(0.132)	(0.130)	(0.126)	(0.197)
Age of household head squared	0.000***	0.000***	0.000***	0.000
	(0.000)	(0.000)	(0.000)	(0.000)
Literacy (dummy)	0.254***	0.229***	0.053	0.062*
	(0.036)	(0.036)	(0.037)	(0.037)
Asset value (log)	–0.015***	–0.002	–0.001	0.000
	(0.005)	(0.005)	(0.005)	(0.005)
Credit (dummy)	0.033	0.025	–0.028	–0.029
	(0.040)	(0.039)	(0.039)	(0.039)
Extension (dummy)	0.047	–0.004	–0.057	–0.060
	(0.034)	(0.034)	(0.050)	(0.050)
Total farm size (ha)	0.155***	0.103***	0.108***	0.071**
	(0.031)	(0.028)	(0.028)	(0.029)
Farm size squared	–0.009**	–0.006*	–0.005*	–0.004
	(0.004)	(0.003)	(0.003)	(0.003)
Rainfall (log)	–0.024***	–0.023***	–0.010	0.028
	(0.008)	(0.008)	(0.008)	(0.017)
North (dummy)		–0.430***	–0.414***	–0.331***
		(0.059)	(0.060)	(0.067)
South (dummy)		–0.931***	–0.967***	–0.961***
		(0.036)	(0.036)	(0.036)
2013 (dummy)			1.103***	1.114***
			(0.059)	(0.060)
2016 (dummy)			0.321***	0.343***
			(0.070)	(0.076)
2019 (dummy)			1.056***	1.084***
			(0.069)	(0.082)
Age of the household head (mean)				0.008
				(0.007)
Asset value (mean)				0.000
				(0.000)
Total farm size (mean)				0.045***
				(0.016)
Rainfall (mean)				–0.000**
				(0.000)
Constant	–0.327	0.554	–0.039	0.322
	(0.515)	(0.507)	(0.493)	(0.623)
Wald □2	183***	848***	1417***	1424***
Observations (No. of households)	6948	6948	6948	6948

Note: Coefficient estimates from RE and Mundlak regressions are shown with robust standard in parentheses. * p < 0·10, ** p < 0·05, *** p < 0·01.

Table A11. Effects of subsidizing legume seeds on productivity: control function approach

	All legumes
=1 if HH received subsidized legume seeds	8.831**
	(4,194)
First stage residual	–7.143**
	(4.177)
Household characteristics	Yes
Region fixed effects	Yes
Year fixed effects	Yes
Mundlak effects	Yes
R-squared	0.119
Observations (No. of households)	6928

Note: To account for zero values for productivity, we transformed our dependent variable (total value of production for legumes (MK/ha)) using IHS transformation ($\log(x+(x^2+1)^{0.5})$). Coefficient estimates from Mundlak regressions are shown with bootstrapped standard errors in parentheses. Other covariates were included for estimation but are not shown here for brevity. ** $p < 0{\cdot}05$.

Table A12. Effects of subsidizing legume seeds on farm production diversity: full results for Figure 3

	Farm production diversity (Crops+Livestock)		Farm production diversity (Crops only)		Farm production diversity (Livestock only)	
	RE	Mundlak	RE	Mundlak	RE	Mundlak
	(1)	(2)	(3)	(4)	(5)	(6)
=1 if HH received subsidized legume seeds	1.292***	1.272***	1.296***	1.276***	1.292***	1.270***
	(0.043)	(0.040)	(0.049)	(0.047)	(0.056)	(0.052)
Male household head	1.058*	1.054*	1.015	1.010	1.223***	1.220***
	(0.025)	(0.025)	(0.025)	(0.025)	(0.051)	(0.051)
Age of the household head (log)	1.233**	1.178	1.088	1.101	1.840***	1.514*
	(0.084)	(0.112)	(0.079)	(0.113)	(0.233)	(0.252)
Age of household head squared	1.000	1.000	1.000	1.000	1.000**	1.000**
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Literacy (dummy)	0.953*	0.967	0.952*	0.967	0.942	0.954
	(0.023)	(0.022)	(0.023)	(0.022)	(0.039)	(0.039)
Asset value (log)	1.009*	1.012***	0.996	0.999	1.057***	1.060***
	(0.003)	(0.003)	(0.003)	(0.003)	(0.008)	(0.008)
Credit (dummy)	1.103***	1.101***	1.112***	1.108***	1.078*	1.080*
	(0.026)	(0.026)	(0.031)	(0.030)	(0.039)	(0.039)
Extension (dummy)	1.035	1.029	0.985	0.980	1.246***	1.240***
	(0.032)	(0.031)	(0.033)	(0.032)	(0.067)	(0.066)
Total farm size (ha)	1.194***	1.139***	1.210***	1.159***	1.149***	1.087***
	(0.028)	(0.019)	(0.038)	(0.028)	(0.021)	(0.015)
Farm size squared	0.995*	0.996*	0.995	0.996	0.997*	0.998*
	(0.002)	(0.001)	(0.003)	(0.002)	(0.001)	(0.001)
Rainfall (log)	1.012*	1.034*	1.008	1.037*	1.026**	1.033
	(0.005)	(0.015)	(0.005)	(0.018)	(0.009)	(0.019)
North (dummy)	1.052	1.106	0.927	0.988	1.335**	1.361**
	(0.079)	(0.081)	(0.088)	(0.092)	(0.122)	(0.128)
South (dummy)	1.248***	1.254***	1.393***	1.397***	0.883	0.891
	(0.066)	(0.064)	(0.085)	(0.084)	(0.057)	(0.055)
2013 (dummy)	2.077***	2.062***	2.115***	2.090***	1.951***	1.965***
	(0.048)	(0.050)	(0.056)	(0.056)	(0.055)	(0.061)
2016 (dummy)	1.136***	1.135**	1.251***	1.238***	0.823***	0.843**
	(0.043)	(0.044)	(0.054)	(0.054)	(0.047)	(0.051)
2019 (dummy)	1.467***	1.455***	1.730***	1.688***	0.847*	0.878
	(0.066)	(0.074)	(0.085)	(0.092)	(0.056)	(0.063)
Mundlak (time averages) test (□2)		52.200***		37.140***		45.780***
Age of the household head (mean)		1.002		1.000		1.009
		(0.003)		(0.003)		(0.005)
Asset value (mean)		1.000*		1.000**		1.000
		(0.000)		(0.000)		(0.000)
Total farm size (mean)		1.066***		1.060***		1.081***
		(0.011)		(0.012)		(0.014)
Rainfall (mean)		1.000		1.000		1.000
		(0.000)		(0.000)		(0.000)
Observations (No. of households)	6907	6907	6907	6907	6907	6907

Note: Coefficient estimates from Poisson regressions are shown with robust standard errors clustered at enumeration area in parentheses. * $p < 0{\cdot}10$, ** $p < 0{\cdot}05$, *** $p < 0{\cdot}01$.

Table A13. Effects of subsidizing legume seeds on farm production diversity: control function approach

	Farm production diversity (Crops+Livestock)	Farm production diversity (Crops only)	Farm production diversity (Livestock only)
	(1)	(2)	(3)
=1 if HH received subsidized legume seeds	3.948***	2.973***	7.847***
	(1.239)	(0.932)	(5.178)
First stage residual	–0.321***	–0.428***	–0.161***
	(0.101)	(0.134)	(0.106)
Household characteristics	Yes	Yes	Yes
Region and year fixed effects	Yes	Yes	Yes
Mundlak effects	Yes	Yes	Yes
Pseudo-R^2	0.170	0.155	0.118
Observations (No. of households)	6907	6907	6907

Note: Coefficient estimates from Poisson regression models are shown with robust standard errors clustered at enumeration area in parentheses. Other covariates were included for estimation but are not shown here for brevity. *** p < 0·01.

Table A14. Effects of subsidizing legume seeds on dietary diversity, calorie and micronutrient consumption: full model results for Table 4

	HDDS	FCS	Calories	Vitamin A	Iron	Zinc
	(1)	(2)	(3)	(4)	(5)	(6)
=1 if HH received subsidized legume seeds	1.045***	0.066**	0.095***	0.099*	0.021	0.066*
	(0.010)	(0.029)	(0.030)	(0.056)	(0.040)	(0.039)
Male household head	0.994	0.002	–0.278***	–0.330***	–0.300***	–0.300***
	(0.007)	(0.026)	(0.022)	(0.032)	(0.022)	(0.023)
Age of the household head (log)	0.987	0.011	–0.850***	–0.616***	–0.692***	–0.903***
	(0.031)	(0.155)	(0.099)	(0.164)	(0.094)	(0.098)
Age of household head squared	1.000*	0.000	0.000***	0.000***	0.000***	0.000***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Literacy (dummy)	1.071***	0.191***	0.135***	0.172***	0.113***	0.161***
	(0.007)	(0.028)	(0.019)	(0.028)	(0.019)	(0.017)
Asset value (log)	1.020***	0.047***	0.025***	0.030***	0.017***	0.022***
	(0.001)	(0.004)	(0.003)	(0.004)	(0.003)	(0.003)
Credit (dummy)	1.022***	0.028	0.029	–0.021	0.009	0.002
	(0.006)	(0.020)	(0.018)	(0.026)	(0.019)	(0.018)
Extension (dummy)	1.001	–0.020	0.043	–0.020	–0.012	–0.010
	(0.009)	(0.016)	(0.027)	(0.040)	(0.027)	(0.028)
Total farm size (ha)	1.000	0.005	0.010	0.000	–0.006	–0.001
	(0.002)	(0.012)	(0.007)	(0.010)	(0.009)	(0.008)
Farm size squared	1.000	0.001*	0.000	0.000	0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Rainfall (log)	0.996	0.007	–0.017*	–0.023	–0.007	–0.001
	(0.003)	(0.009)	(0.009)	(0.018)	(0.010)	(0.011)
North (dummy)	1.025	–0.007	–0.142**	–0.181	–0.112*	–0.103*
	(0.019)	(0.053)	(0.059)	(0.112)	(0.065)	(0.059)
South (dummy)	1.013	–0.033	0.067**	–0.024	0.036	–0.008
	(0.014)	(0.034)	(0.034)	(0.055)	(0.034)	(0.038)
2013 (dummy)	1.000	1.376***	0.062	0.683***	0.511***	0.697***
	(0.009)	(0.073)	(0.038)	(0.060)	(0.044)	(0.043)
2016 (dummy)	0.935***	0.683***	–0.226***	0.129*	–0.070	0.040
	(0.013)	(0.072)	(0.038)	(0.066)	(0.049)	(0.046)
2019 (dummy)	0.971*	0.771***	–0.015	0.672***	0.460***	0.617***
	(0.013)	(0.081)	(0.041)	(0.072)	(0.048)	(0.051)
Mundlak joint test (□2)	6.970	11.320**	31.800***	6.660	7.820*	19.430***
Age of the household head (mean)	1.000	–0.004	0.013***	0.007	0.011***	0.016***
	(0.001)	(0.006)	(0.003)	(0.006)	(0.004)	(0.004)

Asset value (mean)	1.000*	0.000***	0.000***	0.000***	0.000***	0.000***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Total farm size (mean)	0.999	–0.007	0.001	–0.005	0.003	–0.009
	(0.003)	(0.011)	(0.010)	(0.014)	(0.009)	(0.010)
Rainfall (mean)	1.000	0.000	0.000***	0.000	0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Constant	7.021***	3.226***	11.244***	8.836***	5.400***	5.808***
	(0.720)	(0.449)	(0.335)	(0.520)	(0.295)	(0.323)
Wald □2	713***	2116**	772***	775***	1500***	1867***
Observations (No. of households)	6928	6928	6825	6825	6825	6825

Note: HDDS and FCS, denotes household dietary diversity score and food consumption score, respectively. Calories and micronutrient—e.g., vitamin A (μg RE/day/AE), iron (mg/day/AE) and zinc (mg/day/AE)—consumption are expressed as logarithm. Coefficient estimates from Mundlak regressions are shown with robust standard errors clustered at enumeration area in parentheses. * $p < 0{\cdot}10$, ** $p < 0{\cdot}05$, *** $p < 0{\cdot}01$.

Table A15. Effects of subsidizing legume seeds on dietary diversity and calorie consumption: control function approach

	HDDS	Calories
	(1)	(2)
=1 if HH received subsidized legume seeds	1.461***	1.398***
	(0.183)	(0.352)
First stage residual	–0.714***	–1.307***
	(0.068)	(0.353)
Household characteristics	Yes	Yes
Region fixed effects	Yes	Yes
Year fixed effects	Yes	Yes
Mundlak effects	Yes	Yes
Pseudo R^2/R^2	0.017	0.134
Observations (No. of households)	6928	6825

Note: HDDS, household dietary diversity score. Calories consumption are expressed as logarithm. Coefficient estimates from Mundlak regressions are shown with bootstrapped standard errors in parentheses. Other covariates were included for estimation but are not shown here for brevity. *** p < 0·01.

Table A16. Effects of subsidizing maize seeds and fertilizer on dietary diversity, calorie and micronutrient consumption

	HDDS	FCS	Calories	Vitamin A	Iron	Zinc
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Subsidized maize seed						
=1 if HH received subsidized	1.017**	0.043	0.022	0.079	0.032	0.029
maize seed	(0.008)	(0.033)	(0.027)	(0.052)	(0.030)	(0.029)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Pseudo R^2/R^2	0.017	0.250	0.133	0.154	0.222	0.283
Observations	6948	6948	6845	6845	6845	6845
Panel B: Subsidized fertilizer						
=1 if HH received subsidized	1.001	0.005	0.017	0.049	-0.018	-0.021
fertilizer	(0.008)	(0.028)	(0.022)	(0.041)	(0.024)	(0.026)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Pseudo R^2/R^2	0.017	0.249	0.133	0.154	0.222	0.283
Observations	6948	6948	6845	6845	6845	6845

Note: HDDS and FCS, denotes household dietary diversity score and food consumption score, respectively. Calories and micronutrient—e.g., vitamin A (µg RE/day/AE), iron (mg/day/AE) and zinc (mg/day/AE)—consumption are expressed as logarithm. Coefficient estimates from Mundlak regressions are shown with robust standard errors clustered at enumeration area in parentheses. Other covariates were included for estimation but are not shown here for brevity. ** p < 0·05.

Table A17. Links between subsidizing legume seeds and child WAZ

	RE	RE	RE	RE	Mundlak	Mundlak
	(1)	(2)	(3)	(4)	(5)	(6)
=1 if HH received subsidized legume seeds	0.122**	0.103*	0.101*	0.067	0.088	0.058
	(0.058)	(0.056)	(0.056)	(0.056)	(0.056)	(0.056)
Male child (dummy)		–0.083***	–0.083***	–0.080**	–0.083***	–0.082***
		(0.031)	(0.031)	(0.031)	(0.031)	(0.031)
Age of children`s mother (log)		–0.076**	–0.074**	–0.043	–0.022	0.032
		(0.037)	(0.037)	(0.038)	(0.035)	(0.037)
Age of children`s mother squared		0.000***	0.000***	0.000	0.000***	0.000
		(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Age of children`s father (log)		0.006	0.004	–0.031	0.000	–0.048
		(0.035)	(0.035)	(0.035)	(0.033)	(0.034)
Asset value (log)		0.021***	0.021***	0.024***	0.019***	0.022***
		(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
Credit (dummy)		0.044	0.043	0.047	0.041	0.047
		(0.031)	(0.031)	(0.031)	(0.031)	(0.031)
Extension (dummy)		–0.276***	–0.274***	–0.141***	–0.285***	–0.079**
		(0.030)	(0.030)	(0.040)	(0.030)	(0.040)
Total farm size (ha)		–0.019**	–0.019**	–0.020**	–0.012	–0.013
		(0.008)	(0.008)	(0.008)	(0.012)	(0.012)
North (dummy)			0.093*			0.139***
			(0.054)			(0.054)
South (dummy)			0.005			0.024
			(0.034)			(0.033)
2013 (dummy)				0.101*		–0.007
				(0.057)		(0.057)
2016 (dummy)				–0.128**		–0.285***
				(0.056)		(0.057)
2019 (dummy)				–0.206***		–0.441***
				(0.060)		(0.062)
Mundlak joint test (□2)					142***	169***
Age of a child (mean)					–0.089***	–0.092***
					(0.008)	(0.008)
Asset value (mean)					0.000*	0.000*
					(0.000)	(0.000)
Total farm size (mean)					–0.005	–0.003
					(0.018)	(0.018)
Rainfall (mean)					0.000	–0.000***
					(0.000)	(0.000)
Constant	–0.459***	–0.306***	–0.312***	–0.246***	–0.021	0.250***
	(0.016)	(0.056)	(0.059)	(0.067)	(0.065)	(0.082)
R-squared	0.001	0.022	0.023	0.026	0.047	0.054
Wald □2	4**	147***	151***	180***	271***	346***
Observations (No. of children)	6769	6765	6765	6765	6765	6765

Note: WAZ, weight-for-age Z-scores. Coefficient estimates from RE and Mundlak regressions are shown with robust standard errors clustered at household level in parentheses. * p < 0·10, ** p < 0·05, *** p < 0·01.

Table A18. Links between subsidizing legume seeds and child HAZ

	RE	RE	RE	RE	Mundlak	Mundlak
	(1)	(2)	(3)	(4)	(5)	(6)
=1 if HH received subsidized legume seeds	0.076	0.081	0.070	0.021	0.071	0.011
	(0.071)	(0.070)	(0.070)	(0.070)	(0.069)	(0.070)
Male child (dummy)		–0.179***	–0.180***	–0.178***	–0.178***	–0.179***
		(0.036)	(0.036)	(0.036)	(0.036)	(0.036)
Age of children`s mother (log)		–0.157**	–0.153**	–0.080	–0.063	0.000
		(0.067)	(0.067)	(0.068)	(0.067)	(0.069)
Age of children`s mother squared		0.000	0.000	0.000	0.000	–0.000*
		(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Age of children`s father (log)		0.086	0.082	0.024	0.058	0.011
		(0.065)	(0.065)	(0.065)	(0.064)	(0.065)
Asset value (log)		0.034***	0.032***	0.032***	0.033***	0.029***
		(0.005)	(0.005)	(0.005)	(0.005)	(0.005)
Credit (dummy)		–0.022	–0.022	–0.038	–0.023	–0.039
		(0.036)	(0.036)	(0.036)	(0.036)	(0.036)
Extension (dummy)		–0.044	–0.04	–0.095**	–0.017	–0.064
		(0.037)	(0.037)	(0.047)	(0.037)	(0.048)
Total farm size (ha)		–0.012	–0.009	–0.006	–0.006	0.001
		(0.013)	(0.013)	(0.013)	(0.020)	(0.019)
North (dummy)			0.147**			0.181***
			(0.064)			(0.065)
South (dummy)			0.079**			0.098**
			(0.039)			(0.039)
2013 (dummy)				0.402***		0.347***
				(0.074)		(0.074)
2016 (dummy)				0.422***		0.372***
				(0.073)		(0.074)
2019 (dummy)				0.193**		0.174**
				(0.075)		(0.078)
Mundlak joint test (□2)					123***	101***
Age of a child (mean)					–0.066***	–0.060***
					(0.006)	(0.006)
Asset value (mean)					0.000***	0.000***
					(0.000)	(0.000)
Total farm size (mean)					–0.004	–0.001
					(0.024)	(0.024)
Rainfall (mean)					0.000	0.000
					(0.000)	(0.000)
Constant	–1.014***	–1.017***	–1.062***	–1.224***	–0.865***	–1.083***
	(0.019)	(0.069)	(0.073)	(0.083)	(0.075)	(0.094)
R-squared	0.000	0.016	0.017	0.025	0.032	0.039
Wald □2	1	90***	98***	151***	212***	260***
Observations (No. of children)	7971	7968	7968	7968	7968	7968

Note: HAZ, Height-for-age Z-scores. Coefficient estimates from RE and Mundlak regressions are shown with robust standard errors clustered at household level in parentheses. * p < 0·10, ** p < 0·05, *** p < 0·01.

Table A19. Effects of subsidizing legume seeds on farm productivity (intensive margins)

	Farm production diversity	Area planted to legume crops	Gross value of production (all crops)	Gross value of production (all legume crops)
	(1)	(2)	(3)	(4)
HH received subsidized legume seeds (kg)	1.076***	0.059***	0.128**	0.482***
	(0.014)	(0.018)	(0.051)	(0.122)
Male household head	1.054*	-0.074***	0.091	-0.255
	(0.025)	(0.022)	(0.126)	(0.224)
Age of the household head (log)	1.202	-0.067	-0.998**	-0.472
	(0.117)	(0.094)	(0.493)	(0.799)
Age of household head squared	1.000	0.000*	-0.000**	0.000
	(0.000)	(0.000)	(0.000)	(0.000)
Literacy (dummy)	0.967	0.035	-0.003	0.132
	(0.022)	(0.022)	(0.091)	(0.203)
Asset value (log)	1.011***	0.002	0.032**	0.029
	(0.003)	(0.003)	(0.014)	(0.027)
Credit (dummy)	1.098***	0.007	0.132	0.347**
	(0.026)	(0.020)	(0.091)	(0.163)
Extension (dummy)	1.034	-0.02	0.542***	0.074
	(0.031)	(0.023)	(0.187)	(0.276)
Total farm size (ha)	1.141***	0.170***	0.007	0.535***
	(0.019)	(0.012)	(0.045)	(0.080)
Farm size squared	0.996*	-0.004***	-0.001	-0.015***
	(0.002)	(0.000)	(0.001)	(0.002)
Rainfall (log)	1.033*	0.028**	0.036	-0.004
	(0.015)	(0.011)	(0.085)	(0.121)
North (dummy)	1.112	-0.099	-1.100**	-1.927***
	(0.081)	(0.085)	(0.456)	(0.728)
South (dummy)	1.265***	-0.257***	-0.220	0.483
	(0.064)	(0.043)	(0.231)	(0.467)
2013 (dummy)	2.079***	0.710***	0.902***	1.565***
	(0.050)	(0.041)	(0.121)	(0.277)
2016 (dummy)	1.133**	0.111***	0.876***	0.322
	(0.045)	(0.032)	(0.268)	(0.356)
2019 (dummy)	1.458***	0.396***	1.721***	3.594***
	(0.073)	(0.041)	(0.253)	(0.435)
Age of the household head (mean)	1.002	0.000	0.053***	0.021
	(0.003)	(0.003)	(0.014)	(0.025)
Asset value (mean)	1.000*	0.000	0.000	-0.000**
	(0.000)	(0.000)	(0.000)	(0.000)
Total farm size (mean)	1.067***	0.062***	0.154***	0.412***
	(0.011)	(0.013)	(0.058)	(0.108)
Rainfall (mean)	1.000	-0.000***	-0.001	0.000
	(0.000)	(0.000)	(0.001)	(0.001)
Constant	0.799	0.314	12.956***	4.442
	(0.257)	(0.317)	(1.639)	(2.819)
Wald □2	3232***	853***	407***	655***
Observations (No. of households)	6927	6948	6948	6948

Note: To account for zero values for productivity, we transformed our dependent variable (value of production (MK/ha)) using inverse hyperbolic sine transformation ($\log(x + (x^2 + 1)^{0.5})$). Coefficient estimates from Mundlak regressions are shown with robust standard errors clustered at enumeration area in parentheses. * p < 0·10, ** p < 0·05, *** p < 0·01.

Table A20. Effects of subsidizing legume seeds on dietary diversity, calorie and micronutrient consumption (intensive margins)

	HDDS	FCS	Calories	Vitamin A	Iron	Zinc
	(1)	(2)	(3)	(4)	(5)	(6)
HH received subsidized legume seeds (kg)	1.016***	0.028***	0.027**	0.035*	0.000	0.021
	(0.003)	(0.010)	(0.011)	(0.021)	(0.015)	(0.014)
Male household head	0.994	0.001	-0.278***	-0.330***	-0.300***	-0.300***
	(0.006)	(0.026)	(0.022)	(0.032)	(0.022)	(0.023)
Age of the household head (log)	0.989	0.020	-0.851***	-0.620***	-0.696***	-0.903***
	(0.031)	(0.155)	(0.099)	(0.164)	(0.094)	(0.098)
Age of household head squared	1.000*	0.000	0.000***	0.000***	0.000***	0.000***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Literacy (dummy)	1.071***	0.191***	0.134***	0.172***	0.113***	0.160***
	(0.007)	(0.028)	(0.019)	(0.028)	(0.019)	(0.017)
Asset value (log)	1.020***	0.047***	0.025***	0.030***	0.017***	0.022***
	(0.001)	(0.004)	(0.003)	(0.004)	(0.003)	(0.003)
Credit (dummy)	1.022***	0.028	0.028	-0.022	0.009	0.001
	(0.006)	(0.020)	(0.018)	(0.026)	(0.019)	(0.018)
Extension (dummy)	1.002	-0.019	0.045*	-0.017	-0.011	-0.008
	(0.009)	(0.016)	(0.027)	(0.039)	(0.027)	(0.028)
Total farm size (ha)	1.000	0.005	0.01	0.000	-0.006	-0.001
	(0.002)	(0.012)	(0.007)	(0.010)	(0.009)	(0.008)
Farm size squared	1.000	0.001*	0.000	0.000	0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Rainfall (log)	0.996	0.008	-0.016*	-0.022	-0.006	0.000
	(0.003)	(0.009)	(0.009)	(0.018)	(0.010)	(0.011)
North (dummy)	1.026	-0.006	-0.141**	-0.18	-0.112*	-0.102*
	(0.019)	(0.052)	(0.058)	(0.112)	(0.065)	(0.058)
South (dummy)	1.014	-0.032	0.071**	-0.022	0.038	-0.006
	(0.014)	(0.033)	(0.034)	(0.054)	(0.034)	(0.038)
2013 (dummy)	1.000	1.374***	0.066*	0.685***	0.514***	0.699***
	(0.009)	(0.073)	(0.038)	(0.060)	(0.044)	(0.043)
2016 (dummy)	0.934***	0.681***	-0.225***	0.131**	-0.068	0.041
	(0.013)	(0.071)	(0.038)	(0.067)	(0.050)	(0.047)

2019 (dummy)	0.971*	0.772***	-0.009	0.681***	0.465***	0.623***
	(0.013)	(0.081)	(0.041)	(0.073)	(0.048)	(0.051)
Age of the household head (mean)	1.000	-0.005	0.013***	0.008	0.012***	0.017***
	(0.001)	(0.006)	(0.003)	(0.006)	(0.004)	(0.004)
Asset value (mean)	1.000*	0.000***	0.000***	0.000***	0.000***	0.000***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Total farm size (mean)	0.999	-0.007	0.001	-0.005	0.003	-0.01
	(0.003)	(0.010)	(0.010)	(0.014)	(0.009)	(0.010)
Rainfall (mean)	1.000	0.000	0.000***	0.00	0.00	0.00
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Constant	6.981***	3.199***	11.234***	8.831***	5.401***	5.796***
	(0.715)	(0.448)	(0.336)	(0.521)	(0.295)	(0.324)
Wald □2	708***	2120***	754***	778***	1512***	1858***
Observations (No. of households)	6948	6948	6845	6845	6845	6845

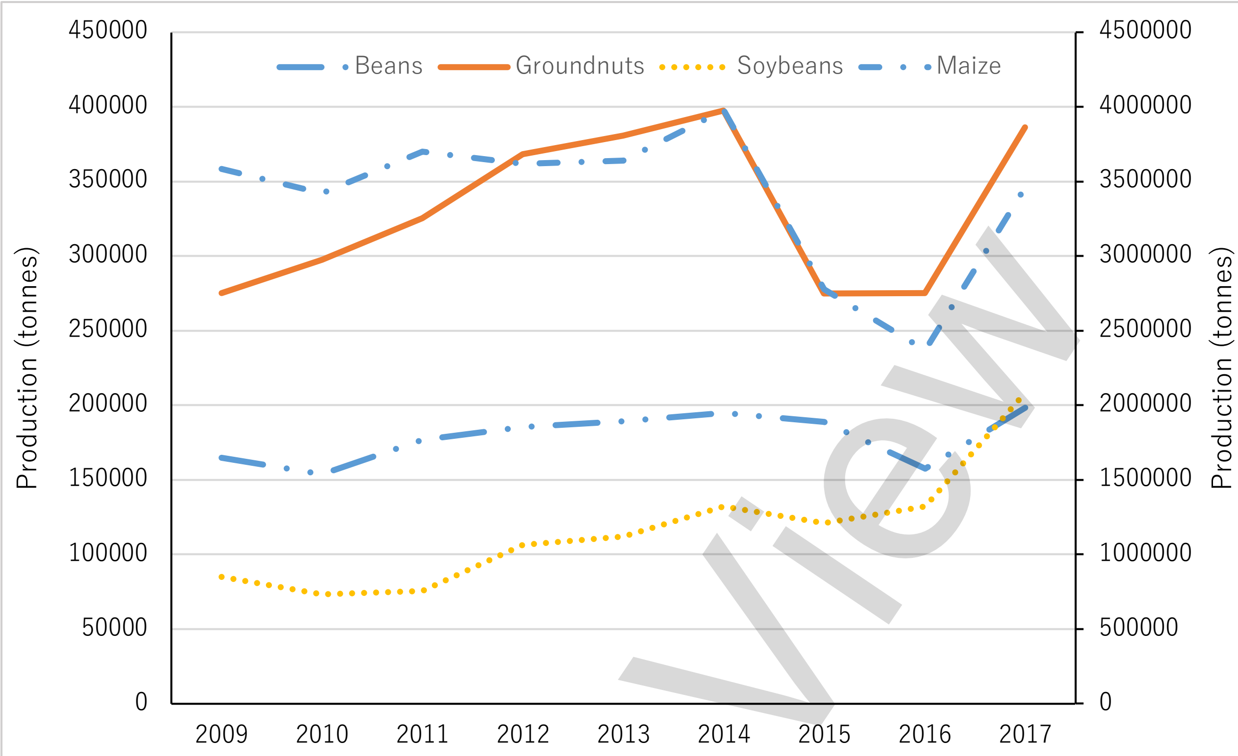
Note: HDDS and FCS, denotes household dietary diversity score and food consumption score, respectively. Calories and micronutrient—e.g., vitamin A (µg RE/day/AE), iron (mg/day/AE) and zinc (mg/day/AE)—consumption are expressed as logarithm. Coefficient estimates from Mundlak regressions are shown with robust standard errors clustered at enumeration area in parentheses. * $p < 0{\cdot}10$, ** $p < 0{\cdot}05$, *** $p < 0{\cdot}01$.

Table A21. Effects of subsidizing legume seeds on farm productivity by farm size category

Dependent variable:	Gross value of production (All crops)			Gross value of production (All legumes)		
Farm size category:	<0.2 ha	0.2-2.5 ha	>2.5 ha	<0.2 ha	0.2-2.5 ha	>2.5 ha
Covariates	(1)	(2)	(3)	(4)	(5)	(6)
=1 if HH received subsidized	3.287***	0.283**	-1.276	4.293***	1.630***	-3.737*
legume seeds	(0.798)	(0.133)	(1.160)	(1.541)	(0.288)	(2.190)
Male household head	-0.164	0.119	0.224	-0.491	-0.263	-0.816
	(0.487)	(0.112)	(0.760)	(0.572)	(0.231)	(0.945)
Age of the household head (log)	-5.057**	-0.042	-5.329**	-1.674	-0.544	-8.665
	(2.438)	(0.430)	(2.574)	(2.808)	(0.843)	(7.801)
Age of household head squared	-0.001*	0.000	0.000	0.000	0.000	0.001
	(0.001)	(0.000)	(0.000)	(0.001)	(0.000)	(0.001)
Literacy (dummy)	-0.565	0.163*	-0.463	-0.462	0.291	-0.096
	(0.486)	(0.084)	(0.330)	(0.567)	(0.200)	(0.969)
Asset value (log)	-0.033	0.058***	-0.101	-0.087	0.045*	-0.142
	(0.064)	(0.013)	(0.108)	(0.073)	(0.027)	(0.193)
Credit (dummy)	0.597	0.032	0.151	0.577	0.253	0.276
	(0.627)	(0.070)	(0.384)	(0.658)	(0.168)	(0.860)
Extension (dummy)	0.243	0.602***	0.429	0.921	-0.026	1.32
	(0.803)	(0.168)	(0.772)	(0.846)	(0.286)	(2.001)
Total farm size (ha)	24.554***	-0.133	0.178	1.409	1.608***	0.303
	(4.701)	(0.120)	(0.135)	(6.440)	(0.260)	(0.242)
Farm size squared	-48.982***	-0.008	-0.003	-7.642	-0.206***	-0.009*
	(10.482)	(0.022)	(0.002)	(18.808)	(0.042)	(0.005)
Rainfall (log)	0.077	-0.172**	-0.496*	-0.004	-0.099	-0.517
	(0.186)	(0.079)	(0.283)	(0.261)	(0.136)	(0.398)
North (dummy)	-1.852*	-0.901***	-2.290***	-0.558	-2.034***	-5.600***
	(1.097)	(0.274)	(0.657)	(0.718)	(0.762)	(1.719)
South (dummy)	0.406	-0.420*	-1.963**	2.682***	0.302	-3.627***
	(0.521)	(0.218)	(0.842)	(0.664)	(0.480)	(1.013)
2013 (dummy)	-2.515***	1.122***	1.671***	-0.195	1.766***	1.411
	(0.582)	(0.093)	(0.573)	(0.750)	(0.300)	(1.245)
2016 (dummy)	1.610*	0.36	1.758	0.482	0.301	0.78
	(0.966)	(0.251)	(1.115)	(1.109)	(0.382)	(2.538)
2019 (dummy)	3.193***	1.089***	1.396	2.443**	3.629***	0.699
	(1.091)	(0.199)	(1.080)	(1.078)	(0.462)	(2.660)
Age of the household head	0.206**	0.020	0.105*	0.052	0.018	0.096
	(0.089)	(0.012)	(0.060)	(0.102)	(0.026)	(0.222)
Asset value (mean)	0.000	0.000	0.000	0.000	-0.000*	-0.000***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Total farm size (mean)	0.472***	0.071	0.090	0.574**	0.303**	0.276
	(0.154)	(0.055)	(0.085)	(0.227)	(0.121)	(0.186)
Rainfall (mean)	-0.002*	0.001**	0.003	-0.001	0.001	0.003
	(0.001)	(0.001)	(0.002)	(0.001)	(0.001)	(0.003)
Constant	24.841***	10.368***	30.019***	8.816	3.991	40.067
	(7.898)	(1.435)	(9.762)	(9.171)	(2.988)	(26.266)
Wald □2	651***	530***	155***	128***	572***	71***
Observations (No. of households)	591	6214	143	591	6214	143

Note: To account for zero values for productivity, we transformed our dependent variable (gross value of production (MK/ha)) using inverse hyperbolic sine transformation ($\log(x + (x^2 + 1)^{0.5})$). Coefficient estimates from RE and Mundlak regressions are shown with robust standard errors clustered at enumeration area in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Figure A1. Production estimates for maize and selected legumes over time (2009–2017) in Malawi



Data source: FAOstat 2020 (http://www.fao.org/faostat/en/#data/QC).

Figure A2. Kernel density distribution of calories and nutrient consumption by beneficiary of subsidized legume seeds