Plant different, eat different? Insights from participatory agricultural research

December 2015

(Please do not cite)

Beliyou Haile, Carlo Azzarri, Cleo Roberts, John Ulimwengu, and Josee Randriamamonjy

This paper examines the link between agricultural production and food consumption using quasi-experimental data from an agricultural research program in Malawi. Using a participatory approach, the program has been testing alternative technology options for sustainable intensification of the maize-pulse systems. Identification is based on comparison of outcomes for program beneficiary and non-beneficiary households. Given the non-random selection into the program and possible simultaneity between production and consumption decisions, the instrumental variables method is used to examine the link. In addition, and considering the program's focus on maize and pulses, a nonlinear food demand system is estimated to assess consumption behavior with respect to different food groups. Program beneficiaries have significantly more diverse agricultural production, relative to non-beneficiaries. While there is an overall positive association between production diversity and dietary diversity in our sample, the improvement in production diversity among program beneficiaries does not translate into more diverse diets. Besides, demand for pulses is the least responsive to expenditure, while that for starchy and animal-source foods is the most responsive to expenditure and prices. These findings highlight the need to rethink crop-specific recommendations in the light of their relative importance in household budget, and suggest that production-oriented agricultural interventions may need to be accompanied by other efforts, such as nutritional education, to maximize their health benefits.

Keywords: production diversity, dietary diversity, food demand, instrumental variables, Malawi

JEL codes: D12, O12, Q18

1. Introduction

Micronutrient deficiency is a widespread problem in developing countries, and the diversity of production into nutrient-rich crops and animal-source foods is among the policy options being pursued to tackle the problem. Adoption of new and improved agricultural technologies that boost production diversity can influence diets—both directly, by improving food availability for subsistence-oriented farmers, and indirectly, by enhancing the purchasing power of commercially oriented farmers. The magnitude of the direct and indirect effects of production diversity is mediated by several factors, such as the level of market integration, the responsiveness of food demand to income, and the nature of the items towards which production is being diversified (for example, whether the items are complements or substitutes to those in the food basket and whether they are staple foods or high-value cash crops).

Where markets are incomplete and transaction costs are high, households with a more diverse production may have better-quality diets (due to the direct effect) than those with a more polarized and subsistence production (Hirvonen and Hoddinott 2014). If diversification is into the production of high-value cash crops, the direct effect could be minimal, with the magnitude of the indirect effect determined by the responsiveness of food demand to income. Rakotoarisoa et al. (2011) note that when households spend a significant portion of income on food, and food consumption is highly income-elastic, income shocks are likely to influence food consumption. On the other hand, if production is diversified into food items whose demand is less responsive to income, the indirect effect of production diversity could be weak. As such, to what extent production diversity affects dietary diversity is an empirical question. Finally, the extent to which food demand and subsequent intake of crucial nutrients respond to changes food price is also important from a policy perspective. For example, households may replace micronutrient-rich (and expensive) diets with high-carbohydrate (and relatively cheaper) staple diets in the face of price inflation.

This paper first examines the causal association between agricultural production and food consumption using cross-sectional data from an agricultural research program in Malawi. The program has been testing various integrated agricultural technologies for maize–pulse systems since 2012. The technologies being tested include improved varieties of maize, groundnut, pigeon pea, cowpea, soybean, and beans, as well as fertilizers. Our identification of causal links is based on comparison of crop-livestock production and food consumption between the two study groups—households testing various mixes of the technologies (hereafter beneficiary households) and randomly sampled non-participating households from non-program villages (hereafter control households). Given the non-random selection of beneficiaries and potential simultaneity between production and consumption decisions, the instrumental variables (IV) approach is used to examine causal links.

Next, a food demand system is estimated based on five good groups (starchy foods, pulses, animal-source foods, fruits and vegetables, and "other" residual food group) to examine possible heterogeneity in the responsiveness of food demand by food group and beneficiary status.[1] Given the nonlinearity of the Engel curves observed in our sample, a quadratic version of the Almost Ideal Demand System (Banks et al. 1997; Deaton and Muellbauer 1980) is estimated, while addressing censoring in budget shares, possible endogeneity in household food expenditure, and heterogeneity in preferences.

We find a significant positive effect of program participation on the diversity and value of agricultural production, thanks to pulse production. While a positive association is found between production diversity and dietary diversity in the study sample, the production-side improvement witnessed by beneficiary households does not translate into a more diverse diet. Pulses are the least responsive to expenditure, while the demand for starchy and animal-source foods (ASF) is the most responsive to expenditure and prices.

The remainder of the paper is organized as follows. A review of the empirical literature on agricultural production, diet, and food demand is presented in Section 2. Study setting, design, and cross-section data used in the empirical analysis are described in Section 3. Section 4 presents a descriptive summary of the data; while Section 5 outlines the identification strategy. Estimation results are presented in Section 6 before concluding the paper in Section 7.

2. Related Literature

Agriculture plays a crucial role in the fight against undernutrition by improving the availability of (nutritious) food and enhancing the income and (food) purchasing power of the poor. Whether and

[1] A list of food items in each food group is summarized in Appendix Table A1.

how pathways from agricultural production to food consumption can lead to improved food and nutrition security depends on several factors. On the food consumption side, these pathways are affected by households' consumption choices over harvest, use of crop income, control over household resources, social status of women, nutritional knowledge, and health and nutritional status of household members (Aberman et al. 2015; Jones et al., 2014; Pinstrup-Andersen 2013). The link between household production and availability of nutritious food may be stronger in settings where markets are weakly integrated and households face challenges in fulfilling their food requirements through purchase. More commercialized households are found to have a higher likelihood of consuming more diverse food than subsistence households (Jones et al. 2014; Sibhatu et al. 2015) and a stronger positive association is found between production diversity and dietary diversity for households with better market access (Hirvonen and Hoddinott 2014). While theoretical literature on the link between production and consumption is well developed, empirical evidence on the relationship between the two is limited (Dillon et al. 2014).

The existing evidence from developing countries suggests positive association between production diversity, dietary diversity, and nutritional outcomes (Azzarri et al. 2015; Bellon et al. 2015; Bhagowalia et al. 2012; Dillon et al. 2014; Hirvonen and Hoddinott 2014; Hoddinott et al. 2015; Jones et al. 2014; Muller 2009). Muller (2009) finds the diversity of crop production, such as beans and tubers, to contribute to improved nutrition among autarkic agricultural households in Rwanda. Dillon et al. (2014) find significant effects of both crop production diversity and agricultural revenue on dietary diversity in Nigeria while a positive association is found between farm diversity and dietary diversity among Malawian households (Jones et al. 2014), with the latter effect being stronger for female-headed and wealthier households. With regard to livestock, Bhagowalia et al. (2012) find ownership of dairy cows (poultry) to increase household milk (meat) consumption among Indian households while Hoddinott et al. (2015) find cow ownership in Ethiopia to increase children's milk consumption. Similarly, Azzarri et al. (2015) find ownership of livestock, especially small ruminants, to improve both household-level ASF consumption and child nutritional outcomes among Ugandan households.

Empirical work has also examined the nutritional effects of specific agricultural technologies. For example, a study from South America finds participation in an agricultural intervention involving several native grain and tuber crops to increase crop diversity and the quantity of program target-crops consumed by participating households (Bellon et al. 2015). While Smale et al. (2015) document a positive association between hybrid seed use and dietary diversity in Zambia, Shiferaw et al. (2014) find higher food consumption and reduced vulnerability to shocks for Ethiopian households adopting improved wheat varieties, relative to non-adopters.

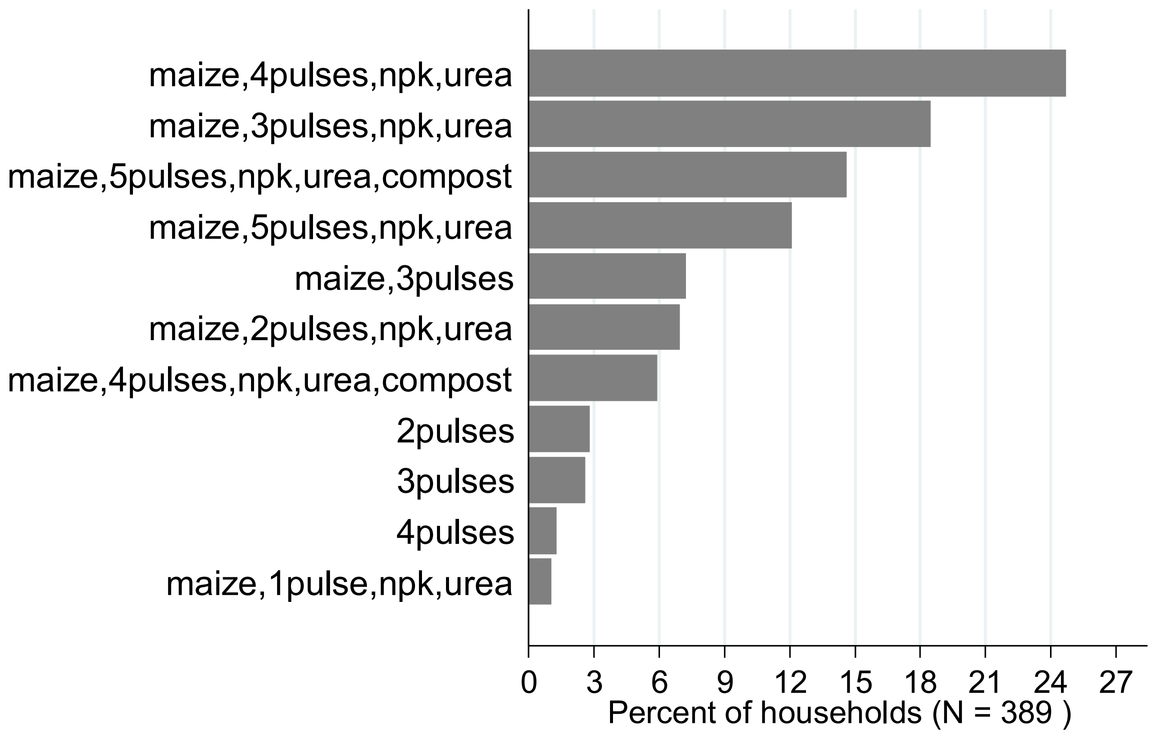
Regarding the responsiveness of food and nutrient demand to prices and income, recent empirical findings from Africa south of the Sahara provide useful insights, despite differences in food grouping and time periods (for example, Boysen (2012) for Uganda; Ecker and Qaim (2011) for Malawi; Abdulai and Aubert (2004) for Tanzania). Ecker and Qaim (2011) find a positive, though declining, association between expenditure elasticities and income for all food groups considered except sugar and sweets in rural areas and fats and oils in urban areas. Boysen (2012) notes a similar exception for sugar, fats, and oils as well as fruit and vegetables, and finds ASF to be luxury items. The same study finds that all food groups considered—except ASF—are price inelastic for rural residents, which could be indicative of their subsistence orientation. For Malawi, Ecker and Qaim (2011) report fats and oils as well as sugar and sweets to be luxuries for rural residents, and ASF to be luxury items for urban dwellers. Unlike Boysen's (2012) findings for Uganda, Ecker and Qaim's (2011) study shows that the demand for all food groups except cassava (in rural areas) and regular beans (in urban areas) is highly price responsive. Abdulai and Aubert (2004) find higher expenditure elasticities of the demand for ASF and micronutrients, relative to cereals and pules.

3. Setting, Study Design, and Data

The focus of this study is a participatory agricultural research program in Malawi called Africa RISING. The program employs "mother and baby trials" to test and identify a menu of technologies to sustainably intensify maize–legume systems in Malawi. These are adaptive research platforms where lead ("mother") farmers actively participate in interactive, researcher-designed, scientifically replicable demonstration trials. Mother trials are established around farmers action groups, whose members subsequently set up "baby" trials to test a subset (or a composition) of the mother-tested technologies on their plots. As of June 2013, more than 400 baby farmers, drawn from 26 villages in Dedza and Ntcheu districts in the Central Region of Malawi, participated in the program. Before program target villages were identified by program researchers, the focus districts were stratified using elevation and temperature-adjusted historical average rainfall.

Beneficiary (baby) farmers, purposively selected by the program implementers,[2] tested agricultural technologies that include maize, a variety of pulses, and fertilizer. A farm-level summary in Figure 3.1 shows the technologies that are tested vary from just two pulses to nine different technologies involving maize, nitrogen-phosphorus-potassium (NPK) compound, urea, compost, and five pulses (groundnut, pigeon pea, cowpea, soybean, and beans).

Figure 3.1 Farm-level technologies tested by program beneficiaries



Source: Africa RISING, Malawi.
Note: Technology mixes tested by less than 1 percent of households not shown.
Pulses refer to the following crops: groundnut, pigeon pea, cowpea, soybean, and beans.
Npk refers to nitrogen-phosphorus-potassium compound. N means number of households

A plot-level summary of these technologies shows groundnut, fertilized maize, and soybean to be the three most common technologies, followed by pigeon pea, cowpea, and fertilized maize intercropped with a pulse (cowpea, beans, or pigeon pea). Intercropping of two pulses, such as

[2] According to an unpublished project document, a baby farmer shall select no more than four (integrated) mother-tested technologies and a plot devoted to each (integrated) technology should be at least 10 meters by 10 meters.

pigeon pea intercropped with soybean or groundnut, is also common. Previous studies have shown maize–pulse intercropping to improve soil fertility, yield, and nutrition while simultaneously reducing fertilizer requirements (Friesen and Palmer 2004; Rajendran et al. 2014). Pulses have better quality (and higher) protein than most plant-based foods and are an important source of dietary fiber, folate, calcium, and iron.

The identification of causal association is based on the comparison of production and food consumption outcomes for program beneficiary households and control households randomly drawn from 28 project non-target villages.[3] The non-target sites are chosen such that they have similar elevation and rainfall (proxies of the agroecology) to those of program target villages, while distant enough from the latter to minimize possible contamination. A household and community survey (Malawi Africa RISING Baseline Evaluation Survey) was conducted in the summer of 2013 (IFPRI, 2015), three to six months after beneficiaries collected their first harvest since joining the program. The survey included detailed crop-level information about agricultural inputs and harvest for the October 2012 to May 2013 season, as well as food consumed (purchased, own-produced, or received as gifts) in the household based on a seven-day recall. The analysis herein is based on a sample of 935 households, of which 397 are beneficiary and 538 are control.

Three variables are constructed to measure household dietary quality (measured by diversity) and quantity (measured by monetary value). The first is a simple count of unique crop and ASF items consumed. The second is annualized total household food expenditure, computed as the sum of expenditure on food purchases and the imputed value of food consumed out of own production or received as gifts.[4] For the third outcome variable, and following Ecker and Qaim (2011), five broad food groups are defined—starchy food, pulses, ASF, fruits and vegetables, and other food—based on their nutritional properties. Next, the Simpson's dietary diversity index (DDI) is constructed (Simpson 1949). This index of richness and evenness is given by $1 - \sum_{j=1}^{5} \left(e_j / e_T\right)^2$, where j indexes food group, e_j is household expenditure on j (in Malawi Kwacha—MWK), and e_T is the

[3] Beneficiaries were selected by project implementers while control households were selected using multistate sampling — identification of four control sites (sections), selection of target villages using probability proportional to size, and random sampling of a fixed number of households (20) per control village.
[4] Unit value (as shadow price) was used to estimate the value of own-produced food and food received as gifts from the associated quantities. Unit value was computed as the ratio between total value and quantity purchased (based on Deaton 1988), where missing unit value has been replaced by the median unit value for the same item for progressively increasing administrative units (village, section, extension planning area, and district).

value of total household food expenditure. Simpson's index ranges between zero and one, with values increasing with the number of food groups consumed, the evenness of the budget share distributions, or both. Household dietary diversity is found to be a good proxy for nutrient adequacy and food security (Hoddinott and Yohannes 2002; Swindale and Bilinsky 2006).

On the production side, three variables mimicking the food consumption variables have been constructed—the count of unique crop and ASF items produced, the value of harvest (as proxy for income), and Simpson's agricultural production diversity index (PDI). The Simpson's PDI is based on the same five groups as for food consumption, constructed using the share of each group of total value of agricultural production. Value of production is expressed by the sum of value of harvest, self-reported value of livestock slaughtered and consumed, and value of animal by-products (meat, milk, and eggs). Self-reported sale prices have been used to value harvest. To control for the effects of household wealth and market access, two indexes are constructed using factor analysis (principal-component factor method), following Filmer and Pritchett (2001). The first is an asset-based wealth index based on a series of questions asked about housing condition and ownership of various durable non-agricultural assets. The second is an access to basic services index based on self-reported travel time to various services.[5] Household- and food group–level aggregate prices have been computed as weighted average of the price of individual food items.[6]

4. Descriptive Summary

Table 4.1 summarizes variables and significance levels from equality of mean tests between beneficiary and control households. Beneficiary households have bigger household size and land area and are more likely to be male headed, relative to control households.

---

[5] These services include motorable road, all-season road, asphalt road, market (daily and weekly), district capital, bus stop, healthcare facility, and schools.

[6] Each food item's price is weighted by its share of expenditure within each food group so that food items consumed in small quantities do not exert undue influence on the aggregate price of the food group. For each food group j, weighted price (p_j) is computed as $\sum_{f=1}^{N} p_{fj} * \left(e_{fj}/e_j\right)$, where e_{fj} is the expenditure share of food item f in group j, and e_j is the expenditure share of food group j of the total household food expenditure. For a household that did not consume any of the items in a food group, the village median of the aggregate price for the food group was used.

Table 4.1 Summary statistics by beneficiary status

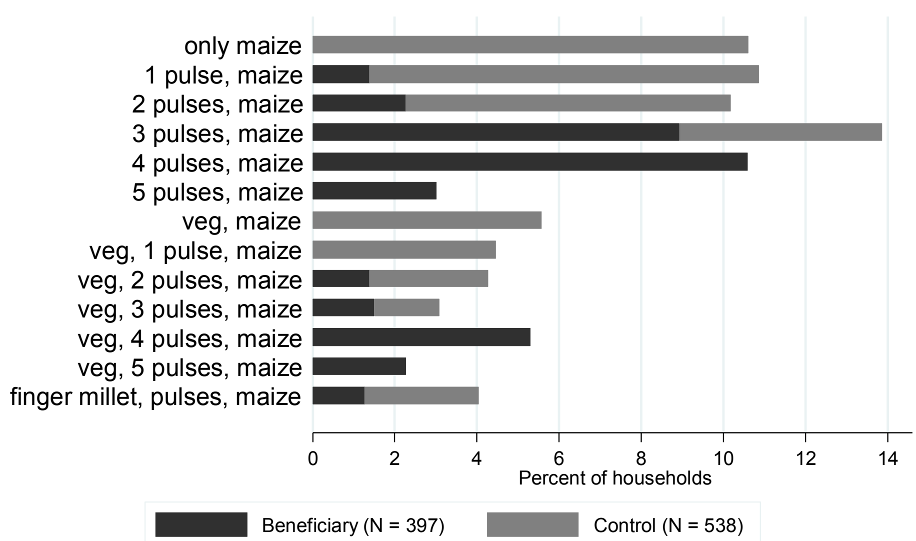
	Beneficiary	Control	Stat. sign
Household size	4.97	4.59	***
Household head age (years)	45.8	45.3	
Female head (%)	0.27	0.34	**
Number of adults (age>=15)	2.66	2.45	***
Number of children (age<15)	2.30	2.13	*
Has a female with at least eight years of schooling (%)	0.13	0.15	
Area of parcels operated (hectare)	1.20	0.86	***
Has a nearby parcel (%)	0.74	0.54	***
Altitude (meters)	864.6	945.6	***
Number of intercropped plots	1.88	1.16	***
Received extension service last year (%)	0.92	0.41	***
Travel time to seed supplier (minutes)	45.4	41.5	
Distance to basic services index	0.036	-0.044	
Own at least one ax (%)	0.64	0.59	*
Nonagricultural wealth index	0.14	-0.068	***
Agricultural labor used (person-days)	332.1	227.9	***
Total fertilizers applied (kilogram)	120.9	80.0	***
Count of crop and animal-based food items produced	6.24	3.98	***
Total harvest value ('000 MWK)	213.8	123.7	***
Harvest value of program-target crops ('000 MWK)	138.9	107.6	***
Harvest value of pulses ('000 MWK)	41.4	30.2	***
Harvest value of starchy foods ('000 MWK)	124.5	88.7	***
Simpson production diversity index	0.40	0.30	***
Count of crop- and animal-based food items consumed	10.7	9.83	***
Simpson dietary diversity index	0.57	0.58	
Annual per capita food expenditure ('000 MWK)	63.4	52.3	***
Experienced food shortage (June-Sep 2013)	0.073	0.18	***
Zero expenditure on starchy foods (%)	0.018	0.017	
Zero expenditure on fruits and vegetables (%)	0.045	0.058	
Zero expenditure on animal source foods (%)	0.29	0.35	*
Zero expenditure on pulses (%)	0.15	0.20	**
Zero expenditure on other foods (%)	0.053	0.082	*
Expenditure share of starchy foods	0.47	0.44	**
Expenditure share of fruits and vegetables	0.12	0.13	*
Expenditure share of animal source foods	0.16	0.17	
Expenditure share of pulses	0.13	0.13	
Expenditure share of other foods	0.12	0.13	
Aggregate price of starchy foods	339.2	435.0	***
Aggregate price of fruits and vegetables	135.9	162.6	**
Aggregate price of animal source foods	1055.3	1035.7	
Aggregate price of pulses	598.8	542.2	*
Aggregate price of other foods	522.6	472.9	**
Observations	397	538	

Source: IFPRI, 2015.

Beneficiaries are also more likely to have a parcel within 15 minutes of the homestead (using usual transport), intercropped plots, received advice about agricultural and resource management practices in the previous year, and reside in a lower-elevation area. They are also better endowed with non-agricultural wealth and report using more agricultural inputs (fertilizers and labor) during the reference period. Both the diversity of agricultural production and the value of harvest are higher for beneficiaries, relative to control households. Specifically, the simple count of crop and ASF items produced and Simpson's PDI are 6.24 and 0.40 for beneficiaries, respectively, while the corresponding values for control households are 3.98 and 0.30. The value of total harvest and harvest of program-target crops—maize, groundnut, pigeon pea, cowpea, and soybean—are about 213,000 MWK and 139,000 MWK for beneficiaries and 124,000 MWK and 107,000 MWK for control households, respectively. Although the difference is not statistically significant, beneficiaries appear to reside in slightly more remote areas, as can be seen from the positive score on the distance to basic services index.

Figure 4.1 shows that beneficiaries are more likely to combine maize with several pulses, while control households are more likely to grow only maize or combine maize with either vegetables or fewer pulses.

Figure 4.1 Farming systems by beneficiary status

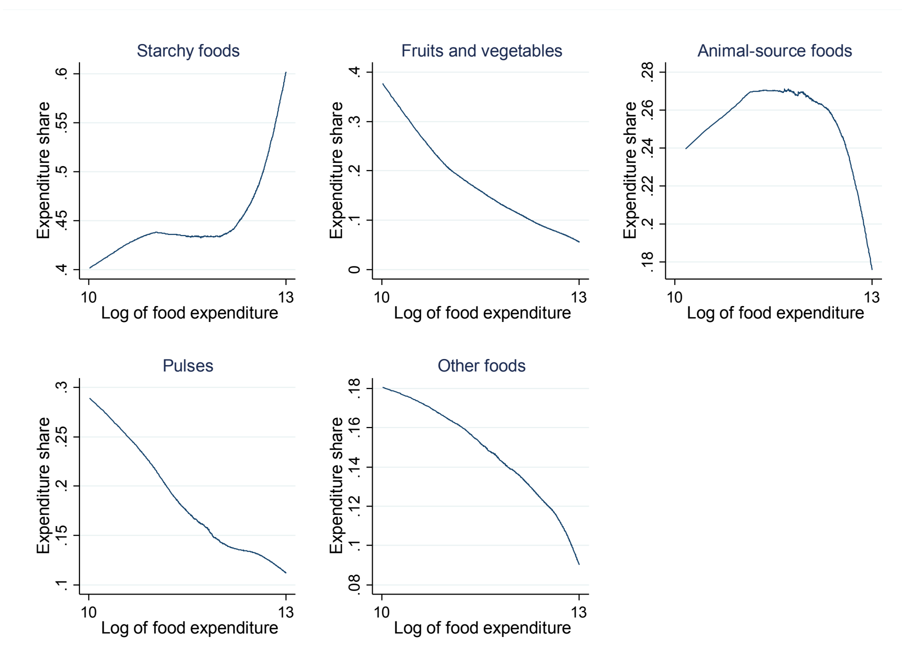


Source: IFPRI, 2015.
Note: Crop mixes grown by less than 2 percent of households not shown.
Pulses refer to the following crops: groundnut, pigeon pea, cowpea, soybean, and beans.
Veg means vegetables.
N means number of households

On the food consumption side, beneficiaries show a marginally higher number of crops and ASF items consumed (10.7 versus 9.83 for control households) and per capita annual food expenditure (63,000 MWK versus 52,000 MWK for control households), with statistically similar Simpson's DDI of 0.58 for the two groups (Table 4.1). The incidence of zero food expenditure varies from about 2 percent on starchy foods to 35 percent on ASF. Estimation of the food demand model employs the two-step approach (Shonkwiler and Yen 1999)[7] to address this censoring of food budget shares. Almost half of the food expenditure is on starchy foods, with control households more likely to pay a higher price for starchy foods as well as fruits and vegetables (although the latter are the least expensive). Beneficiary households are less likely to report facing food shortage between June and September 2013.

To guide the choice of food demand functional form, a locally weighted regression (Cleveland and Devlin 1988) of expenditure shares is estimated. A summary of Engel curves in Figure 4.2 shows nonlinearity for the food groups considered, justifying a nonlinear model.

Figure 4.2 Nonparametric estimation of Engel curves



Source: IFPRI, 2015.

[7] See the Appendix for details.

Among the socioeconomic variables considered that could affect food consumption choices, budget shares appear to vary by household size and elevation of the household's residence, as summarized in Appendix Figure A1. The budget share equations therefore controls for these two variables.

5. Identification

5.1 Program participation, Agricultural Production, and Food Consumption

The following system of equations is estimated to examine causal links between program participation, agricultural production, and food consumption:

$$Prod_i = \alpha_0 + \alpha_1 Treat_i + \mathbf{\Phi}'\boldsymbol{X}_{1i} + d + \epsilon_i \quad (1)$$
$$Cons_i = \beta_0 + \beta_1 Prod_i + \beta_2 Treat_i + \beta_3 (TreatxProd_i) + \mathbf{B}'\boldsymbol{X}_{2i} + d + \mu_i \quad (2)$$

where i indexes household. Equation 1 models agricultural production ($Prod$) through one of the variables discussed in Section 3; the potentially endogenous variable $Treat$ is an indicator that takes a value of 1 for beneficiary household and zero otherwise; $\boldsymbol{X}_1$ is a vector of covariates that could affect diversity and value of agricultural production and includes household size, age and gender of the head, index of distance to basic services, non-agricultural wealth index, total land area operated, indicator for ownership of at least one ax, agricultural labor and fertilizers used, and travel time to the nearest agricultural input supplier; d represent district fixed effect; and ϵ expresses a random error term. If the participation decision is not random (but correlated with ϵ), $Treat$ will be endogenous and ordinary least squares (OLS) estimation of Equation 1 would produce biased and inconsistent estimates.

Equation 2 models food consumption ($Cons$) using one of the variables defined in Section 3. The conditioning variables $Prod$, $Treat$, and $TreatxProd$ are potentially endogenous, with the last one capturing the expected difference in the slope estimate of $Prod$ between beneficiary and control households; $\boldsymbol{X}_2$ is a vector of covariates that could affect household food consumption and consists of all the variables in $\boldsymbol{X}_1$ (except agricultural inputs) and an indicator of whether the household faced food shortage between June and September 2013; μ is a random error term. If the production and consumption decisions are nonseparable or omitted variables affect both, $Prod$ will be endogenous, as will $Treat$ if the participation decision is not random with respect to μ.

The estimate $\widehat{\alpha_1}$ measures the likely effect of program participation on production. When the dependent variable in Equation 2 is the count of crops and ASF items consumed (Simpson's DDI), $\widehat{\beta_1}$ measures the effect of a unit increase in the number of crops and ASF items produced (Simpson's PDI) for the control group, while $\widehat{\beta_3}$ measures the difference in slope between the two groups. A statistically significant and positive $\widehat{\beta_3}$ suggests a unit increase in each of the production variables to have a stronger effect for beneficiaries than for control households and vice versa. When regressing food expenditure on harvest value, both expressed in natural logarithm, $\widehat{\beta_1}$ and $\widehat{\beta_3}$ become elasticities.

To address possible endogeneity, and as a robustness check, three estimators are employed: IV via two-stage least squares (IV/2SLS) (Wooldridge 2013); IV via the generalized method of moments (IV/GMM) (Baum et al. 2007; Hansen 1982); and three-stage least squares (3SLS) (Zellner and Theil 1962). When the assumption of independently and identically distributed (i.i.d) errors is violated and the model is overidentified, IV/GMM generates more efficient estimates than IV/2SLS due to the use of an optimal weighting matrix (Baum et al. 2007, 2003). If cross-equation correlations are due, for example, to nonseparability and there are no misspecifications, both IV/2SLS and 3SLS estimators are consistent but the latter is asymptotically more efficient. If an equation in the system is misspecified, on the other hand, the full-information 3SLS estimator carries the effects of misspecification in any single equation to another part of the system, as it relies on incorrectly estimated variance-covariance matrix from the second stage to recover system parameters (Baltagi 2002). Given the clustering in our sample, heteroskedasticity- and cluster-robust (for the IV/2SLS estimator) and bootstrapped errors (for the 3SLS estimator) (Ando and Hodoshima 2007) are reported.

Since the OLS estimator is consistent and more efficient than the IV estimator in the absence of endogeneity, we first conduct a Hausman test to determine whether the suspect endogenous regressor(s) can be considered exogenous (Hausman 1978). The IV estimator can result in inconsistencies and finite-sample biases when the instrument(s) are weakly correlated with the endogenous variable(s) (Bound et al. 1995; Sargan 1958; Staiger and Stock 1997). Thus, diagnostic tests of instrument relevance are conducted based on the significance of the excluded instruments in the first-stage reduced form regressions (Cragg and Donald 1993; Kleibergen and Paap 2006; Stock and Yogo 2005; Stock et al. 2002). With multiple endogenous regressors, as is the case with the food consumption model, the search for valid instruments gets more difficult since the excluded instruments need to capture enough of the exogenous variation in the endogenous regressors to avoid perfect collinearity in the population regression of the structural equation. Finally, the IV estimator will be inconsistent if the instruments are correlated with the error in the structural equation. Thus, a test of overidentifying restrictions is conducted for the (overidentified) food consumption model using the Sargan (Chi-squared) and Hansen (J) tests (Hansen 1982; Sargan 1958).

The search for relevant and exogenous instruments is guided by the design of the agricultural program of focus, where beneficiaries have been purposefully selected. As summarized in Table 4.1, ownership of a parcel close to the homestead is positively correlated with the likelihood of selection into the program but, we argue, is less correlated with production diversity and whole-farm production (which also includes faraway parcels). The observed correlation could be due to the need for having a nearby parcel for ease of monitoring of the trials by the program researchers and agricultural extension agents. As shown in Section 6, the Hausman test justifies the use of an IV estimator for the production model, except when the dependent variable is harvest value. While we cannot perform an exogeneity test for the exactly identified production model, we cannot reject the null that ownership of a nearby parcel is strongly correlated with program participation based on the econometrics tests reported in Section 6.

For the three potentially endogenous regressors in the food consumption model (program participation, production outcomes, and an interaction term between the two), the following four instruments are used: ownership of a nearby parcel, elevation of household's residence, number of intercropped plots, and whether the household received extension services in the preceding year. Elevation affects agricultural potential and is a proxy for temperature and evapotranspiration, while the intensity of intercropping and likelihood of receiving an extension service are expected to be strongly correlated with agricultural production. Since these instruments capture different aspects of the farming systems of study households and are generally weakly correlated with each other,[8] we expect them to be good instruments. As discussed in Section 6, an endogeneity test justifies the use of the IV estimator, and various diagnostic tests do not reject the null that these

[8] The pairwise correlation coefficients for this set of instruments are –0.05, 0.06, 0.07, 0.16, 0.17 and 0.29.

four excluded instruments are strongly correlated with the suspect endogenous variables and are uncorrelated with the error term of the food consumption model.

5.2 Food Demand

Two potential pathways for agricultural production to affect food consumption are (1) through direct consumption of diverse own-produced food (direct channel), and (2) through higher income due to enhanced cash-crop production, allowing more (diverse) food purchases on the market (indirect channel). The identification in Section 4 does not allow for disentangling the effect of the two channels, but assessing the distribution of expenditure across food groups can shed further light on the relative importance of each channel. For policymaking purposes too, assessing the responsiveness of food demand for different commodities could guide decision-making about whether and how much should be invested on those commodities if the intent is also to promote food security and improve dietary quality.

In light of the nonlinearity in the Engel curves summarized in Figure 4.2, a Quadratic Almost Ideal Demand System (QUAIDS) is estimated (Banks et al. 1997). Our estimation of food demand functions accounts for censoring of budget shares (Shonkwiler and Yen 1999), heterogeneity of preferences (Blow 2003; Pollak and Wales 1981) and possible endogeneity of food expenditure in the budget share equations due to measurement error (Hausman et al. 1995; Lewbel 1996). Demand equations (in budget share forms) are estimated using the iterative feasible generalized nonlinear least squares estimator (Poi, 2012). After recovering the parameters of the expenditure share equations, expenditure and uncompensated (Marshallian) price elasticities are computed at sample means for the total sample as well as for the beneficiary and control groups, the latter to examine possible differential responsiveness between the two groups. See the Appendix for details on the derivation and estimation of the budget share equations.

6. Results and Discussion

6.1 Program participation and Agricultural Production

Regression results on the effect of program participation on diversity and value of agricultural production are presented in Table 6.1.[9] Formal econometric tests of endogeneity and relevance are also provided. Hausman test rejects that program participation is exogenous in all the models, except when harvest value is the dependent variable (Table 6.1, Column 6), justifying IV estimator in the former case but OLS in the latter. The cluster-robust Kleibergen-Paap (KP) F-statistic is around 11, above Staiger and Stock's (1997) rule of thumb value of 10 for a single endogenous regressor, below which the null of weak instrument cannot be rejected. The Cragg-Donald F-statistic, valid under i.i.d error (Baum et al. 2007), is 23, above the Stock-Yogo's (2005) 5 percent critical F of 16.38 with a maximal size distortion of 10 percent. Based on these tests, weak correlation between ownership of a nearby parcel and program participation can be rejected. The predictive power of the Simpson's PDI model is relatively weak, as can be seen from the R-squared in Columns 3 and 4.[10] First-stage reduced-form regression results are reported in Appendix Table A2.

The IV/2SLS estimates show a positive effect of participation on the count of crops and ASF items produced (Table 6.1, Column 2) as well as on Simpson's PDI (Table 6.1, Column 4).[11] Indeed, beneficiaries produced about 4.5 units more and their Simpson's score is 0.35 units higher, relative to control households. Participation positively affect harvest value, with value being 31 percentage points higher for beneficiaries relative to control households (Table 6.1, Column 5).[12] As would be expected, variable inputs (fertilizers and agricultural labor) and fixed agricultural capital (land size and ownership of an ax) are positively correlated with agricultural production.

[9] For the food consumption model, only IV/2SLS results are reported as the IV/GMM estimator reduces to the IV/2SLS estimator for an exactly identified equation (Baum et al. 2003).

[10] Unlike OLS estimation of a linear model with a constant, the R-squared (R2) for IV/2SLS is not constrained between zero and one (inclusive). IV/2SLS is not nested within a constant-only model of the regressand, and IV residuals are computed over a different set of regressors. A negative R2 suggests weak predictive power of the model.

[11] The IV/2SLS estimates are quite large compared with the corresponding OLS estimates (about 2.6 times more for the count variable and 5 times more for the Simpson's PDI), suggesting possible underestimation with the latter estimator.

[12] When the dependent variable is the count of crops and ASF items produced, an exponential conditional mean model (with additive errors) with an endogenous regressor is also estimated via GMM. The difference in the logs of expected counts is 1.07 units higher for beneficiaries, relative to the control group.

Table 6.1 Program participation and agricultural production

	Count of crops and ASF produced				Simpson's production diversity index				Harvest value (in log)			
	1:OLS		2:IV/2SLS		3:OLS		4:IV/2SLS		5:OLS		6:IV/2SLS	
	coef	se	coef	se	coef	se	coef	se	coef	se	coef	se
Participation in the program	1.715***	0.216	4.543***	1.231	0.071***	0.026	0.345***	0.127	0.312**	0.159	0.888	0.756
Household size	0.087*	0.052	0.054	0.057	0.001	0.004	-0.002	0.005	0.009	0.025	0.002	0.024
Age of household head (years)	0.003	0.005	0.002	0.006	0.000	0.000	-0.000	0.000	-0.001	0.003	-0.002	0.003
Max adult yrs of education	0.017	0.025	0.020	0.036	0.004*	0.002	0.005	0.004	0.039**	0.016	0.040**	0.017
Female head	-0.080	0.133	0.020	0.170	-0.040**	0.016	-0.030*	0.016	0.012	0.081	0.033	0.086
Number of adults (age>=15)	-0.023	0.088	0.005	0.117	-0.018**	0.008	-0.015	0.010	0.044	0.054	0.050	0.057
Distance to basic services index	0.068	0.089	0.115	0.150	0.021***	0.008	0.026*	0.014	-0.018	0.053	-0.008	0.067
Area of parcels operated (ha)	0.645***	0.152	0.310	0.274	0.052***	0.012	0.019	0.021	0.279***	0.089	0.211	0.131
Ownes of at least one ax	0.648***	0.128	0.782***	0.179	0.051***	0.013	0.064***	0.018	0.281***	0.074	0.309***	0.089
Non-agr. wealth index	0.070	0.080	0.031	0.096	-0.003	0.007	-0.006	0.009	-0.011	0.041	-0.019	0.042
Agricultural labor used (person-days)	0.001***	0.000	0.000	0.001	0.000**	0.000	0.000	0.000	0.001***	0.000	0.000**	0.000
Total fertilizers applied (kg)	0.002***	0.001	0.002**	0.001	0.000	0.000	0.000	0.000	0.002***	0.000	0.002***	0.000
Travel time to seed supplier (min)	-0.001	0.001	0.000	0.002	-0.000	0.000	0.000	0.000	-0.000	0.001	0.000	0.001
Constant	2.001***	0.338	1.444***	0.504	0.241***	0.034	0.187***	0.049	9.946***	0.215	9.832***	0.288
Number of observations	935		935		935		935		935		935	
R-squared (R2) (Centered)	0.443		0.107		0.171		-0.261		0.380		0.325	
Uncentered R2				0.840				0.684				0.993
Hausman endogeneity test				7.162				5.391				0.803
Hausman P-value				0.007				0.020				0.370
Kleibergen-Paap rk Wald F statistic				11.243				11.243				11.243
Cragg-Donald Wald F statistic				23.259				23.259				23.259
F test of overall model fit		38.058		40.762		6.187		6.196		15.949		12.834
F test P-value				0.000				0.000				0.000

Note: *** p<0.01, ** p<0.05, * p<0.1

District fixed effects included. Reported are cluster-robust standard errors. OLS = Ordinary least squares, IV/2SLS = Two-step instrumental variables.

Endogenous (instrumented) variables in columns 2, 4 and 6: participation in the program.

Excluded instrument in columns 2, 4 and 6: ownership of a nearby parcel.

6.2 Program participation, Agricultural Production, and Food Consumption

The results summarized in Table 6.1 are to be expected, given the (more valuable and exportable) pulses beneficiaries were exposed as part of the program, as summarized in Figure 3.1 and Figure 4.1. While these early stage results are encouraging developments, what also matters from a policy perspective is whether and how much of these gains translate into better-quality diets. As noted before, the impact pathway could be direct (consumption of pulses that are high in fiber, protein, and other micronutrients that are crucial for a healthy and balanced diet) or indirect (consumption of other nutritious and more expensive foods, such as ASF, in light of the enhanced agricultural income) or both.[13]

To explore this further in the context of the program we are focusing on, Table 6.2 summarizes regression results from our food consumption model. The Hausman test rejects that participation, agricultural production (diversity or value), and the interaction term between the two are exogenous. The KP Lagrange multiplier test rejects the null of underidentification while the KP first-stage F-statistic is less than 10, which is expected given the number of endogenous regressors.[14] The Sargan-Hansen overidentification tests do not reject the null that the excluded instruments are uncorrelated with the error term in the food consumption model. The IV/GMM estimator seems to improve the precision of some coefficient estimates but has not changed the qualitative result. The discussion below therefore focuses only on IV/GMM estimates.

A unit increase in the number of own-produced food items increases the number of food items consumed in the household by 1.3 (Table 6.2, Column 3) while a unit increase in the Simpson PDI increases the Simpson's DDI score by 0.37 (Table 6.2, Column 6). On the other hand, harvest value, the proxy measure of income, does not have a significant effect on food expenditure (Table 6.2, Column 9), suggesting possibly weak indirect channel. Coefficients of the interaction terms between participation dummy and each of the production measures (count of items produced in Column 3, Simpson's PDI in Column 6, and harvest value in Column 9) are either insignificant or marginally significant (at 10 percent), suggesting a similar effect of production diversity on dietary diversity for beneficiary and control households.

[13] On the other hand, this direct effect could be minimal if the diversity is into export-oriented food crops. According to a recent press release, for example, Malawi is one of the East Africa nations aiming to meet 4 million metric tons of pulses demanded in India (Nation Online, 2013), a move that could come at the expense of consumption of own-produced pulses.
[14] The first-stage results are summarized in Appendix Table A3.

While it is possible for beneficiary households to have consumed more diverse (and more valuable) food outside the reference week, it is equally likely, we argue, for control households to also have consumed more diverse (and more valuable) food. We do not expect systematically different consumption patterns for the two groups during the reference week. Our results differ from those of Bellon et al. (2015), who find a positive effect of participation in an agricultural program on the quantity of program target-crops consumed. Dietary diversity is negatively associated with household size, age of the household head, access to basic services, and the likelihood of experiencing food shortage between June and September 2013, while better educated households are more likely to consume more diverse food (Table 6.2, Column 3).

Table 6.2 Program participation, production, and food consumption

	Count of food items consumed						Simpson dietary diversity index					
	1:OLS		2:IV/2SLS		3:IV/GMM		4:OLS		5:IV/2SLS		6:IV/GMM	
	coef	se	coef	se	coef	se	coef	se	coef	se	coef	se
Participation in the program (Treat)	0.531	1.054	7.287	6.883	7.533*	4.546	0.008	0.029	0.440*	0.246	0.443*	0.248
Count of crops and ASF items produced (pdi)	0.503***	0.110	1.236**	0.594	1.261***	0.359						
Treatxpdi	-0.209	0.177	-1.186	1.232	-1.236	0.791						
Household size	-0.303**	0.135	-0.320*	0.175	-0.324**	0.155	-0.000	0.005	0.001	0.006	0.001	0.006
Age of household head (years)	-0.039***	0.012	-0.041***	0.013	-0.041***	0.011	-0.001	0.001	-0.001	0.001	-0.001*	0.000
Max adult yrs of education	0.214***	0.060	0.178**	0.073	0.180***	0.060	0.001	0.002	0.002	0.002	0.002	0.002
Female head	-0.172	0.338	-0.040	0.375	-0.059	0.335	-0.001	0.015	-0.005	0.017	-0.005	0.015
Number of children (age<15)	0.147	0.179	0.108	0.209	0.117	0.193	-0.000	0.007	-0.002	0.007	-0.003	0.008
Has a female with at least eight yrs of schooling	0.051	0.401	0.350	0.481	0.357	0.510	-0.000	0.013	0.004	0.016	0.004	0.019
Distance to basic services index	-0.474***	0.168	-0.492***	0.173	-0.488***	0.171	-0.005	0.007	-0.006	0.008	-0.006	0.007
Area of parcels operated (ha)	0.211	0.210	-0.467	0.347	-0.464	0.330	0.015*	0.008	0.008	0.010	0.007	0.012
Ownes of at least one ax	0.850***	0.313	0.880**	0.358	0.910**	0.360	0.031**	0.013	0.040**	0.016	0.040**	0.016
Non-agr. wealth index	0.847***	0.274	0.774***	0.260	0.768***	0.207	0.003	0.009	-0.003	0.009	-0.003	0.008
Experienced food shortage (June-Sep 2013)	-2.107***	0.348	-1.235**	0.542	-1.215**	0.491	-0.041***	0.015	-0.015	0.021	-0.015	0.021
Simpson production diversity index (simpson_pd)							0.129***	0.036	0.376***	0.104	0.378***	0.103
Treatxsimpson_pd							-0.085	0.054	-1.098*	0.593	-1.103*	0.619
Constant	9.149***	0.921	6.339***	2.278	6.222***	1.590	0.567***	0.036	0.495***	0.052	0.495***	0.043
Number of observations	935		935		935		935		935		935	
R-squared (R2) (Centered)	0.239		0.148				0.101		-0.264			
Uncentered R2				0.846						0.912		
Hausman endogeneity test				20.948						10.015		
Hausman P-value				0.000						0.018		
Kleibergen-Paap rk LM statistic				12.594						7.287		
Kleibergen-Paap P-value				0.002						0.026		
Kleibergen-Paap rk Wald F statistic				4.129						2.534		
Cragg-Donald Wald F statistic				7.860		1.394				1.923		0.071
Sargan-Hansen overidentification test				0.837						0.059		
Sargan's P-value				0.360						0.808		
F test of overall model fit				23.083						5.400		
F test P-value		27.386		0.000				6.102		0.000		
Hansen's P-value						0.238						0.789

Note: *** p<0.01, ** p<0.05, * p<0.1. District fixed effects included. Reported are cluster-robust standard errors. OLS = Ordinary least squares, IV/2SLS = Two-step instrumental variables (IV), IV/GMM = IV via Generalized Method of Moments GMM). Endogenous (instrumented) variables in columns 2 and 3: participation in the program (Treat), Count of crop and ASF items produced (pdi) and Treatxpdi. Endogenous (instrumented) variables in columns 5 and 6: Treat, Simpson's production diversity index (simpson_pd) and Treatxsimpson_pd. Excluded instruments in columns 2, 3, 5, and 6: ownership of a nearby parcel, altitude, number of intercropped plots, and received extension service (last 12 months).

Table 6.2 Program participation, production, and food consumption (cont'd)

	Ln annual food expenditure (MWK)					
	m7:OLS		m8:IV/2SLS		m9:IV/GMM	
	coef	se	coef	se	coef	se
Participation in the program (Treat)	-1.204*	0.641	-3.973	4.661	-4.415	4.063
Household size	0.035	0.025	0.016	0.030	0.016	0.027
Age of household head (years)	-0.007***	0.002	-0.007***	0.002	-0.007***	0.002
Max adult yrs of education	0.030***	0.010	0.017	0.012	0.017	0.012
Female head	-0.057	0.061	-0.051	0.070	-0.050	0.060
Number of children (age<15)	-0.009	0.031	-0.005	0.033	-0.003	0.034
Has a female with at least eight yrs of schooling	0.027	0.078	0.135	0.088	0.146	0.093
Distance to basic services index	0.006	0.038	0.014	0.043	0.016	0.032
Area of parcels operated (ha)	0.082*	0.045	-0.031	0.085	-0.029	0.055
Ownes of at least one ax	0.154***	0.059	0.132**	0.060	0.132*	0.068
Non-agr. wealth index	0.136***	0.024	0.108***	0.029	0.107***	0.029
Experienced food shortage (June-Sep 2013)	-0.527***	0.087	-0.367***	0.121	-0.374***	0.098
Ln harvest value (lnval)	0.011	0.041	0.101	0.229	0.090	0.118
Treatxlnval	0.112*	0.057	0.381	0.401	0.419	0.348
Constant	11.693***	0.447	10.860***	2.458	10.989***	1.336
Number of observations	935		935		935	
R-squared (R2) (Centered)	0.249		0.139			
Uncentered R2				0.996		
Hausman endogeneity test				21.992		
Hausman P-value				0.000		
Kleibergen-Paap rk LM statistic				10.588		
Kleibergen-Paap P-value				0.005		
Kleibergen-Paap rk Wald F statistic				3.470		
Cragg-Donald Wald F statistic				4.811		0.677
Sargan-Hansen overidentification test				0.483		
Sargan's P-value				0.487		
F test of overall model fit				19.272		
F test P-value		26.172		0.000		
Hansen's P-value						0.410

Note: *** p<0.01, ** p<0.05, * p<0.1

District fixed effects included. Reported are cluster-robust standard errors. OLS = Ordinary least squares, IV/2SLS = Two-step instrumental variables (IV), IV/GMM = IV via Generalized Method of Moments GMM). Endogenous (instrumented) variables in columns 8 and 9: Participation in the program (Treat), Ln harvest value (lnval) and Treatxlnval. Excluded instruments in columns 8 and 9: ownership of a nearby parcel, altitude, number of intercropped plots, and received extension service (last 12 months).

Finally, 3SLS estimation results are summarized in Table 6.3. The qualitative results generally reinforce those from the single-equation IV/2SLS estimator—namely, participation has a positive effect on the diversity and value of agricultural production, but participants have not enjoyed a more diverse (or valuable) diet.

Table 6.3 Program participation, production, and food consumption (three-stage least squares)

	System 1				System 2				System 3			
	1:pdi		2:ddi		3:simpson_pd		4:simpson_dd		5:lnval		6:lnfexp	
	coef	se	coef	se	coef	se	coef	se	coef	se	coef	se
Participation in the program (Treat)	2.545***	0.254	4.093	4.461	0.052*	0.027	0.455***	0.150	0.560***	0.108	-1.658	3.183
Household size	0.077*	0.041	-0.360**	0.145	0.001	0.004	-0.001	0.006	0.006	0.022	0.016	0.030
Household head age (years)	0.003	0.005	-0.042***	0.012	0.000	0.000	-0.001	0.001	-0.001	0.003	-0.007***	0.002
Max adult yrs of education	0.018	0.021	0.186***	0.053	0.004*	0.002	0.002	0.002	0.040***	0.013	0.019	0.012
Female head (%)	-0.051	0.137	-0.069	0.333	-0.040***	0.014	-0.011	0.015	0.020	0.076	-0.052	0.060
Number of adults (age>=15)	-0.013	0.073			-0.018**	0.008			0.047	0.042		
Distance to basic services index	0.082	0.067	-0.499***	0.164	0.021***	0.007	-0.003	0.007	-0.012	0.037	0.013	0.030
Area of parcels operated (hectare)	0.553***	0.160	-0.512	0.451	0.054***	0.013	0.016	0.017	0.252***	0.063	-0.041	0.070
Ownes of at least one ax	0.689***	0.138	0.695	0.464	0.050***	0.014	0.047***	0.016	0.294***	0.070	0.129*	0.075
Non-agricultural wealth index	0.056	0.084	0.742***	0.267	-0.002	0.007	-0.003	0.008	-0.014	0.037	0.113***	0.026
Agricultural labor used (person-days)	0.001**	0.000			0.000*	0.000			0.000***	0.000		
Total fertilizers applied (kilogram)	0.002***	0.001			0.000	0.000			0.002***	0.000		
Travel time to seed supplier (minutes)	-0.000	0.001			-0.000	0.000			-0.000	0.001		
Count of crops and ASF items produced (pdi)			1.207**	0.480								
Treatxpdi			-0.704	0.785								
Number of children (age<15)			0.129	0.178			-0.000	0.008			-0.001	0.037
Has a female with at least eight years of schooling			0.294	0.502			0.003	0.018			0.106	0.086
Experienced food shortage (June-Sep 2013)			-1.374***	0.492			-0.015	0.019			-0.412***	0.089
Simpson production diversity index (simpson_pd)							0.248	0.160				
Treatxsimpson_pd							-1.116***	0.366				
Ln harvest value (lnval)											0.161	0.171
Treatxlnval											0.180	0.276
Constant	1.835***	0.255	6.885***	1.803	0.244***	0.026	0.526***	0.051	9.899***	0.171	10.153***	1.841
Number of observations	935				935				935			
R-squared	0.414		0.171		0.169		-0.343		0.370		0.169	
Chi-squared	644.980		302.443		165.346		64.149		562.538		321.794	
P-value	0.000		0.000		0.000		0.000		0.000		0.000	
Log-Likelihood	-4,551.96				566.79				-2,342.30			

Note: *** p<0.01, ** p<0.05, * p<0.1

District fixed effect included. Reported are bootstraped standard errors (100 replications). ddi = count of food items consumed, simpson_dd = Simpson's dietary diversity index, lnfexp = ln of annual food expenditure (MWK), pdi = count of crop and ASF items produced, simpson_pd = Simpson's production diversity index, and lnval = ln of harvest value. Each system of equation (pdi and ddi in Systems 1, simpson_pd and simpson_dd in Systems 2, and lnval and lnfexp in Systems 3) is estimated using 3SLS. Endogenous (instrumented) variables in columns 1, 3, and 5 is participation in the program (Treat). Endogenous variables are Treat, pdi and Treatxpdi (column 2), Treat, simpson_pd and Treatxsimpson_pd (column 4), and Treat, lnval and Treatxlnval (column 6). Excluded instrument in columns 1, 3, and 5 is ownership of a nearby parcel. Excluded instruments in columns 2, 4, and 6 are ownership of a nearby parcel, altitude, number of intercropped plots, and received extension service (last 12 months).

Similar to the IV/2SLS estimator, the R-squared from 3SLS is negative for the Simpson's DDI model (Table 6.3, Column 4). Estimates of the effect of participation on agricultural production from IV/2SLS and 3SLS decline, respectively, from 4.5 to 2.5 (when the dependent variable is count of items produced) and from 0.34 to 0.05 (when the dependent variable is Simpson's PDI). Also, the effect of participation on harvest value increases from 31 percent (IV/2SLS) to 56 percent (3SLS). While the magnitude (and significance) of estimates of the effect of production diversity on dietary diversity are comparable for the count dependent variable, effect becomes insignificant when looking at Simpson's DDI. These differences in IV/2SLS and 3SLS estimates may suggest possible misspecifications in either one of the equations in the system that gets carried over to all the system parameters when using the latter estimator.

6.3 Food Demand

Table 6.4 presents expenditure elasticities for the whole sample and separately for beneficiaries and controls. All elasticities are positive and significantly different from zero. ASF are the most (and pulses, the least) responsive to expenditure, with estimates for pulses measured relatively less precisely. As ASF are the most expensive (highest valued) food group, with a price per kilogram almost twice as much as the next-most-expensive food group, their budget share seems to be the most sensitive to changes in total food expenditures. Specifically, a 1 percent reduction in food expenditure reduces the demand for ASF and pulses by 1.12 and 0.59 percent, respectively.

Table 6.4 Expenditure elasticities

	Food groups				
	Starchy foods	Pulses	ASF	Fru/Veg	Other
Total	1.086	0.596	1.124	0.918	1.092
	(0.031)	(0.240)	(0.095)	(0.130)	(0.157)
Beneficiary	1.083	0.568	1.133	0.913	1.096
	(0.030)	(0.259)	(0.104)	(0.139)	(0.166)
Control	1.089	0.615	1.117	0.921	1.089
	(0.032)	(0.228)	(0.088)	(0.125)	(0.152)

Note: Standard errors in parentheses. ASF = Animal source foods. Fru/Veg= Fruits and vegetables. Other = All other foods consumed.

Because expenditures on starchy foods, ASF, and other foods increase at a higher rate than other food expenditures, pulses make up an ever-decreasing share of the budget, suggesting they are a necessity. According to Ecker and Qaim (2011), pulses are highly inelastic among urban Malawian households, and even inferior goods among better-off households, showing an expenditure elasticity close to unity among rural households.[15] It is possible that improved welfare is correlated with higher sales of pulses, given the relatively high pulse price —second only to ASF—and the relative importance of pulses as exportable cash crops in Malawi (Nation Online 2013; Odeny 2007; Simtowe 2000; Snapp et al. 2003).

Expenditure elasticities for the beneficiary and control groups are identical. This implies the absence of a systematic difference in responsiveness irrespective of intergroup differences with regard to not only farm production but also other socioeconomic variables (summarized in Table 4.1) that could affect the responsiveness of food demand to expenditure changes. While one of the arguments for the promotion of pulses (within the program and beyond) is to help improve nutrition (through the production of nutrient-rich crops), findings for pulses do not provide evidence on the existence of this (desirable) link. These results could be interpreted within the context of the findings of Verduzco-Gallo et al. (2014), where per capita pulse consumption for rural Malawian households declined between 2004 and 2011. Verduzco-Gallo et al. (2014) note that the decline has happened in spite of the country's large-scale farm input subsidy program, where pulse seeds were either subsidized or given out for free and where one would expect an improvement in pulse consumption.

Finally, the uncompensated price elasticities are reported in Table 6.5. Consistent with demand theory, all own-price elasticities are negative, with the demand for ASF and pulses being the most and least price elastic, respectively, and no variation in price responsiveness across the two groups. High price elasticity of ASF suggests a possible adverse effect of ASF price inflation on the intake of nutrient-dense ASF. This is especially the case considering Malawi's currently underdeveloped livestock sector and the subsequent weak direct channel between livestock ownership and ASF consumption. Once again, and in spite of the socioeconomic differences between the beneficiary and control households (including market access), price responsiveness is similar for the two groups (Table 6.5, Panels B and C).

[15] Comparison of our expenditure elasticity estimates for pulses may not be accurate, nonetheless, since Ecker and Qaim's (2011) group for pulses includes beans, soybeans, and groundnuts, whereas our grouping has more items as summarized in Appendix Table A1.

Table 6.5 Uncompensated price elasticities

	Panel A. All					Panel B. Beneficiary					Panel C. Control				
	Starchy foods	Pulses	ASF	Fru &Veg	Other	Starchy foods	Pulses	ASF	Fru &Veg	Other	Starchy foods	Pulses	ASF	Fru &Veg	Other
Starchy foods	**-1.007**	-0.077	0.035	0.036	0.032	**-1.006**	-0.080	0.035	0.039	0.033	**-1.007**	-0.075	0.034	0.034	0.032
	(0.026)	(0.076)	(0.022)	(0.041)	(0.047)	(0.025)	(0.083)	(0.025)	(0.045)	(0.050)	(0.027)	(0.071)	(0.020)	(0.038)	(0.045)
Pulses	-0.081	**-0.769**	0.090	-0.021	-0.085	-0.077	**-0.754**	0.096	-0.022	-0.089	-0.083	**-0.780**	0.086	-0.021	-0.082
	(0.022)	(0.091)	(0.023)	(0.035)	(0.048)	(0.021)	(0.097)	(0.026)	(0.038)	(0.051)	(0.023)	(0.086)	(0.021)	(0.033)	(0.046)
ASF	0.020	0.257	**-1.190**	-0.018	0.024	0.019	0.274	**-1.205**	-0.019	0.025	0.021	0.246	**-1.180**	-0.018	0.023
	(0.011)	(0.044)	(0.025)	(0.032)	(0.037)	(0.011)	(0.047)	(0.027)	(0.034)	(0.039)	(0.012)	(0.043)	(0.024)	(0.031)	(0.036)
Fru/Veg	0.001	-0.025	-0.026	**-0.852**	-0.138	0.001	-0.028	-0.028	**-0.845**	-0.144	0.001	-0.024	-0.025	**-0.858**	-0.133
	(0.009)	(0.040)	(0.021)	(0.034)	(0.036)	(0.009)	(0.043)	(0.022)	(0.036)	(0.038)	(0.009)	(0.039)	(0.020)	(0.033)	(0.035)
Other foods	0.002	0.008	0.005	-0.104	**-0.942**	0.002	0.009	0.005	-0.108	**-0.940**	0.002	0.007	0.006	-0.101	**-0.943**
	(0.017)	(0.046)	(0.017)	(0.026)	(0.052)	(0.017)	(0.050)	(0.019)	(0.027)	(0.055)	(0.018)	(0.044)	(0.016)	(0.024)	(0.050)

Note: Standard errors in parentheses. ASF = Animal source foods. Fru/Veg = Fruits and vegetables. Other = All other foods consumed.

7. Conclusions

Production diversity is often considered a promising approach to improve dietary diversity, either by increasing the availability of food to subsistence-oriented smallholders or by enhancing the (food) purchasing power of households, or both. A number of factors mediate the interaction between production and dietary diversity, including market access, the income elasticity of the demand for different food items, and the nature of the crops whose production is diversified. This paper focus on a participatory agricultural program in Malawi to examine the effects of the value and diversity of agricultural production on the diversity and amount of food consumed at the household level.

The program is testing various agricultural technologies aimed at more efficient and profitable maize–pulse farming systems, including fertilized maize, maize–pulse intercropping, and intercropping between two pulses. These technologies are first tested through interactive, researcher–designed, and scientifically replicable demonstration trails. Subsequently, farmers exposed to the technologies through field days select and test a subset of the technologies on their plots. Our identification of the effects of program participation and production diversity is based on the comparison of outcomes for program beneficiary households and randomly selected non-beneficiary (control) households. Given the self-selection into the program and possible simultaneity between production and consumption decisions, the instrumental variables approach is employed to establish causality. In addition, a quadratic food demand model is estimated for different food groups (starchy foods, pulses, animal-source foods, fruits and vegetables) to examine possible differential responsiveness of demand by food group, including those targeted by the program, and beneficiary status.

Beneficiaries show a more diverse agricultural production, relative to the control group, both in terms of simple count of agricultural items produced and group count. Diversity has improved thanks to an enhanced production of pulses – crops that are rich in crucial micronutrients and have better and higher quality protein than other grains. We find an overall positive association between production diversity and dietary diversity that does not vary by beneficiary status, although beneficiaries have witnessed a significant increase in production diversity. On the opposite, we do not find a significant association between harvest value and household food expenditure. The demand for pulses is the least responsive to expenditure, while that of animal-source foods is the most responsive to expenditure, likely due to the progressive role of the former as cash crop, with insignificant difference in expenditure responsiveness between beneficiary and control households.

Our findings may suggest that while improvements in agricultural productivity are necessary to increase access to food and rural household incomes, they are not sufficient to ensure dietary quality and food security. Specifically, agricultural interventions aimed at improving the yield of high-value and nutrient-dense crops could bear limited effects on household consumption of these crops if, for example, knowledge about their nutritional benefits is limited or there are challenges in acquiring more diverse (and nutritious) food from the market. As such, production-oriented interventions may need to go hand in hand with other interventions, such as nutritional education and improvements in the physical access to and the structure of markets.

In our case, program beneficiaries were exposed to the pulse-based technologies mainly through demonstration field days, and they might not have gained enough insights on the health benefits of the target crops or their combination. Indeed, and according to an unpublished document, the program has since then incorporated nutritional education to improve the nutritional knowledge of participants. This could be a crucial step in bridging the observed gap between production diversity (into pulses) and dietary diversity (into pulses). At a global level, 2016 is declared as the International Year of Pulses by the 68th United Nations General Assembly in an effort to promote awareness about the nutritional benefits of pulses and increase their contribution to the fight against food insecurity and undernutrition.

Finally, while our study has provided useful insights on the link between agricultural production and household-level food consumption, it does not address potential intrahousehold inequalities in the distribution of food. Also, food consumption patterns for program beneficiaries may become more responsive to production patterns over time, and especially through the ongoing nutritional education, a limitation of this study we expect to address using follow-up data.

References

Abdulai, A., and D. Aubert. 2004. "A Cross-section Analysis of Household Demand for Food and Nutrients in Tanzania." *Agricultural Economics* 31 (2004): 67–79.

Aberman, N., J. Meerman, and T. Benson. 2015. *Mapping the Linkages between Agriculture, Food Security and Nutrition in Malawi*. IFPRI Malawi Strategy Support Program. Washington, DC: International Food Policy Research Institute.

Ando, M., and J. Hodoshima. 2007. "A Note on Bootstrapped White's Test for Heteroskedasticity in Regression Models." *Economics Letters* 97(1): 46–51.

Azzarri, C., A. Zezza, B. Haile, and E. Cross. 2015. "Does Livestock Ownership Affect Animal Source Foods Consumption and Child Nutritional Status? Evidence from Rural Uganda." *The Journal of Development Studies* 51(8): 1034–1059.

Baltagi, B. H. 2002. *Econometrics,* third ed. New York: Springer.

Banks, J., R. Blundell, and A. Lewbel. 1997. "Quadratic Engel Curves and Consumer Demand." *The Review of Economics and Statistics* 79 (4): 527–539.

Baum, C. F., M. E. Schaffer, and S. Stillman. 2003. "Instrumental Variables and GMM: Estimation and Testing." *The Stata Journal* 3 (1): 1–31.

———. 2007. "Enhanced Routines for Instrumental Variables/GMM Estimation and Testing." *The Stata Journal* 7 (4): 465–506.

Bellon, M. R., E. Gotor, and F. Caracciolo. 2015. "Assessing the Effectiveness of Projects Supporting On-farm Conservation of Native Crops: Evidence from the High Andes of South America." *World Development* 70: 162–176.

Bhagowalia, P., D. Headey, and S. Kadiyala. 2012. *Agriculture, Income, and Nutrition Linkages in India: Insights from a Nationally Representative Survey*. IFPRI Discussion Paper 01195. Washington, DC: International Food Policy Research Institute.

Blow, L. 2003. *Explaining Trends in Household Spending*. Working Paper 03/06. The Institute for Fiscal Studies.

Bound, J., D. Jaeger, and R. Baker. 1995. Problems with Instrumental Variables Estimation When the Correlation between the Instruments and the Endogeneous Explanatory Variable is Weak. Journal of the American Statistical Association, 90(430): 443-450.

Boysen, O. 2012. *A Food Demand System for Uganda*. IIIS Discussion Paper No. 396. Dublin: Institute for International Integration Studies.

Cleveland, W. S., and S. J. Devlin. 1988. "Locally Weighted Regression: An Approach to Regression Analysis by Local Fitting." *Journal of the American Statistical Association* 83 (403): 596–610.

Cragg, J., and S. Donald. 1993. “Testing Identifiability and Specification in Instrumental Variables Models.” *Econometric Theory* 9: 222–240.

Deaton, A. 1988. “Quality, Quantity, and Spatial Variation of Price.” *The American Economic Review* 78 (3): 418–430.

Deaton, A., and J. Muellbauer. 1980. “An Almost Ideal Demand System.” *The American Economic Review* 70 (3): 312–326.

———. 1990. *Price Elasticities from Survey Data: Extensions and Indonesian Results.* Living Standards and Measurement Study Working Paper No. 69. Washington, DC: World Bank.

Dillon, A., K. McGee, and G. Oseni. 2014. *Agricultural Production, Dietary Diversity, and Climate Variability*. World Bank Policy Research Working Paper 7022. Washington, DC: World Bank.

Ecker, O., and M. Qaim. 2011. “Analyzing Nutritional Impacts of Policies: An Empirical Study for Malawi.” *World Development* 39 (3): 412–428.

Filmer, D. and L. Pritchett. 2001. “Estimating Wealth Effects without Expenditure Data—or Tears: An Application to Educational Enrollments in States of India.” *Demography* 38 (1): 115–132.

Friesen, D., and A. Palmer, eds. 2004. *Integrated Approaches to Higher Maize Productivity in the New Millennium*. *Proceedings of the Seventh Eastern and Southern Africa Regional Maize Conference*. CIMMYT (International Maize and Wheat Improvement Center), Kenya Agricultural Research Institute, Nairobi, Kenya, February 5–11.

Hansen, L. 1982. “Large Sample Properties of Generalized Method of Moments Estimators.” *Econometrica* 50: 1029–1054.

Hausman, J. A., W. K. Newey, and J. L. Powell. 1995. “Nonlinear Errors in Variables: Estimation of Some Engel Curves.” *Journal of Econometrics* 65 (1995): 205–233.

Hausman, J. 1978. “Specification Tests in Econometrics.” *Econometrica* 46 (6): 1251–1271.

Hirvonen, K., and J. Hoddinott. 2014. *Agricultural Production and Children’s Diets: Evidence from Rural Ethiopia*. Ethiopia Strategy Support Program Working Paper No. 69. Addis Ababa: International Food Policy Research Institute.

Hoddinott, J., D. Headey, and M. Dereje. 2015. “Cows, Missing Milk Markets, and Nutrition in Rural Ethiopia.” *The Journal of Development Studies* 51 (8): 958–975.

Hoddinott, J., and Y. Yohannes. 2002. *Dietary Diversity as a Food Security Indicator*. Food Consumption and Nutrition Division Discussion Paper 136. Washington, DC: International Food Policy Research Institute.

IFPRI. 2015. "Malawi Africa Research in Sustainable Intensification for the Next Generation (Africa RISING) Baseline Evaluation Survey", http://dx.doi.org/10.7910/DVN/28557, Harvard Dataverse, V1.

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

Kleibergen, F., and R. Paap. 2006. "Generalized Reduced Rank Tests Using the Singular Value Decomposition." *Journal of Econometrics* 133: 97–126.

Lewbel, A. 1996. "Demand Estimation with Expenditure Measurement Errors on the Left and Right Hand Side." *The Review of Economics and Statistics* 78 (4): 718–725.

Muller, C. 2009. "Do Agricultural Outputs of Partly Autarkic Peasants Affect Their Health and Nutrition? Evidence from Rwanda." *Food Policy* 34 (2): 166–175.

Nation Online. 2013. "India Ready to Buy More Pulses from Malawi." Nation Online. March 26. http://mwnation.com/india-ready-to-buy-more-pulses-from-malawi.

Odeny, D. A. 2007. "The Potential of Pigeon Pea (*Cajanus cajan* (L.) Millsp.) in Africa." *Natural Resources Forum* 31 (4): 297–305.

Pinstrup-Andersen, P. 2013. *Guiding Food System Policies for Better Nutrition*. Background paper for the State of Food and Agriculture. Rome: Food and Agriculture Organization of the United Nations.

Poi, B. 2012. Easy demand-system Estimation with QUAIDS. *The Stata Journal*, 12(3): 433–446.

Pollak, R. A., and T. J. Wales. 1978. "Estimation of Complete Demand Systems from Household Budget Data." *The American Econometric Review* 68 (3): 348–359.

———. 1981. "Demographic Variables in Demand Analysis." *Econometrica* 49 (6): 1533–1551.

Rajendran, S., V. Afari-Sefa, M. Bekunda, L. Dominick, and P. Lukumay. 2014. Does Crop Diversity Contribute to Dietary Diversity? Evidence from Integration of Vegetables into Maize-based Farming Systems in Tanzania. Agricultural Economics Society 88th Annual Conference, AgroParisTech, Paris, France, April 9–11.

Rakotoarisoa, M. A., M. Iafrate, and M. Paschali. 2011. *Why Has Africa Become a Net Food Importer? Explaining African Agricultural and Food Trade Deficits*. Rome: Food and Agriculture Organization of the United Nations.

Sargan, J. 1958. The Estimation of Economic Relationships using Instrumental Variables. Econometrica, 26(3) : 393-415.

Shiferaw, B., M. Kassie, M. Jaleta, and C. Yirga. 2014. "Adoption of Improved Wheat Varieties and Impacts on Household Food Security in Ethiopia." *Food Policy* 44 (2014): 272–284.

Shonkwiler, J. S., and S. T. Yen. 1999. "Two-Step Estimation of a Censored System of Equations." *American Journal of Agricultural Economics* 81 (4): 972–982.

Sibhatu, K. T., V. V. Krishna, and M. Qaim. 2015. "Production Diversity and Dietary Diversity in Smallholder Farm Households. Department of Agricultural Economics and Rural Development." *Proceedings of the National Academy of Sciences of the United States of America* 112 (34): 10657–10662.

Simpson, E. H. 1949. "Measurement of Diversity." *Nature* 688: 163.

Simtowe, F. 2000. "Export Agriculture in Malawi: Alternatives to Tobacco." In *Policy Backgrounds: Project Paper No. 1*, Chapter 2. Washington, DC: International Food Policy Research Institute.

Smale, M., M. Moursi, and E. Birol. 2015. "How Does Adopting Hybrid Maize Affect Dietary Diversity on Family Farms? Micro-evidence from Zambia." *Food Policy* 52 (2015): 44–53.

Snapp, S. S., R. B. Jones, E. M. Minja, J. Rusike, and S. N. Silim. 2003. "Pigeon Pea for Africa: A Versatile Vegetable—and More." *Horticultural Science* 38 (6): 1073–1079.

Staiger, D., and J. H. Stock. 1997. "Instrumental Variables Regression with Weak Instruments." *Econometrica* 65: 557–586.

Stock, J., and M. Yogo. 2005. "Testing for Weak Instruments in Linear IV Regression." In *Identification and Inference for Econometric Models: 23 Essays in Honor of Thomas Rothenberg*, edited by D. Andrews and J. Stock: 80–108. Cambridge: Cambridge University Press.

Stock, J., J. Write, and M. Yogo. 2002. A Survey of Weak Instruments and Weak Identification in Generalized Method of Moments. Journal o*f Business & Economic Statistic.* 20(4): 518-529.

Swindale, A., and P. Bilinsky. 2006. *Household Dietary Diversity Score (HDDS) for Measurement of Household Food Access: Indicator Guide (v.2)*. Washington, DC: FHI 360/FANTA.

Verduzco-Gallo, I., O. Ecker, and K. Pauw. 2014. *Changes in Food and Nutrition Security in Malawi: Analysis of Recent Survey Evidence*. IFPRI Working Paper No. 6. Malawi Strategy Support Program.

Wooldridge, J. M. 2013. *Introductory Econometrics: A Modern Approach*, fifth ed. Ohio: South-Western.

Zellner, A., and H. Theil. 1962. "Three-Stage Least Squares: Simultaneous Estimation of Simultaneous Equations." *Econometrica* 30 (1): 54–78.

Appendix

Quadratic Almost Ideal Demand System (QUAIDS) of Food Demand

This section discusses the QUAIDS model estimated, addressing three challenges often faced when estimating a system of food demand equations: censoring of budget share, endogeneity in income (or expenditure), and preference heterogeneity. Following Banks et al. (1997), an indirect utility function consistent with quadratic budget share can be specified as follows:

$$V(\mathbf{P}, Y) = \left[\left(\frac{\ln Y - \ln a(\mathbf{P})}{b(\mathbf{P})}\right)^{-1} + \lambda(\mathbf{P})\right]^{-1} \tag{1}$$

where Y measures income; $\mathbf{P}$ is a vector of food prices; and $a(\mathbf{P})$, $b(\mathbf{P})$, and $\lambda(\mathbf{P})$ are price aggregator functions defined as follows:

$$\ln a(\mathbf{P}) = \alpha_0 + \sum_{j=1}^{n} \alpha_j \ln p_j + \frac{1}{2}\sum_{j=1}^{n}\sum_{k=1}^{n} \gamma_{jk} \ln p_j \ln p_k \tag{2}$$

$$\ln b(\mathbf{P}) = \prod_{j=1}^{n} {p_j}^{\beta_j} \tag{3}$$

$$\lambda(\mathbf{P}) = \sum_{j=1}^{n} \lambda_j \ln p_j \tag{4}$$

where *j* and *k* are indexes of food group (*j, k* = 1, 2,…, n) and n equals 5 in our case. By substituting Equations 2–4 into Equation 1, exploiting the duality in consumer theory, and applying Roy's identity to the expenditure function, the quadratic budget share equations shown in Banks et al. (1997) can be obtained.

When preferences and consumption choices differ by certain characteristics, it is necessary to specify a budget share equation that controls for those characteristics (Blow 2003; Blundell et al. 1994; De Agostini 2014; Pollak and Wales 1981). This paper employs the linear demographic translating method proposed by Pollak and Wales (1978, 1981) and applied in several demand studies (Abdulai 2002; Akbay et al.

2007; Dhar et al. 2003). A (deterministic) quadratic budget share model that controls for demographic characteristics d_l (∀ l) can be written as follows:

$$w_j = \alpha_{oj} + \sum_{l=1}^{l} \theta_{jl} d_l + \sum_{k=1}^{n} \gamma_{jk} \ln p_k + \beta_j \ln \left[\frac{Y}{a(\boldsymbol{P})}\right] + \frac{\lambda_j}{b(\boldsymbol{P})} \left\{ \ln \left[\frac{Y}{a(\boldsymbol{P})}\right] \right\}^2 \tag{5}$$

where w_j is the food budget share; α_{0j}, θ_{jl}, γ_{jk}, β_j, and λ_j are parameters to be estimated; and other variables are as defined before. Our budget share equations control for the two demographic characteristics summarized in the top panel of Appendix Figure A1 below—household size and elevation of household's residence. The parameter restrictions in Equations 6–8 are imposed to fulfil the adding-up, homogeneity, and Slutsky symmetry restrictions discussed in Deaton and Muellbauer (1980) and Paris and Caracciolo (2014).

$$\sum_{j=1}^{n} \alpha_{0j} = 1, \quad \sum_{j=1}^{n} \beta_j = 0, \quad \sum_{j=1}^{n} \lambda_j = 0, \quad \sum_{j=1}^{n} \gamma_{jk} = 0 \;\; \forall\, k \;, \sum_{j=1}^{n} \theta_{jl} = 0 \;\forall\, l \tag{6}$$

$$\sum_{k=1}^{n} \gamma_{jk} = 0 \tag{7}$$

$$\gamma_{jk} = \gamma_{kj} \quad \forall\, j \neq k \tag{8}$$

The censoring of budget shares summarized in Table 1 in the article can be due to optimal decision by the household, lack of consumption of some food groups during the reference week, or measurement error (Johnson 1989). While conditional elasticities can be estimated based only on observations with nonzero budget shares (Deaton 1990), they will be prone to self-selection bias if censoring is non-random. The two-step approach, first proposed by Shonkwiler and Yen (1999) and applied in several demand studies (Aepli 2014; Su and Yen 2000; Tafere et al. 2010; Tefera et al. 2012), is used in this paper. This method of correcting for self-selection involves estimation of density and cumulative density functions at the first stage to be used as weights (for the probability of positive consumption) and additional regressor, respectively, when estimating the budget share equations at the second stage. The method can be formalized as follows:

$$w_j^* = f\left(\boldsymbol{X}_j, \boldsymbol{\Gamma}_j\right) + \nu_j \tag{9}$$

$$q_j^* = \boldsymbol{Z}_j'\Theta_j + \xi \text{ , } q_j = \begin{cases} 1 \; if - \xi_j < \boldsymbol{Z}_j'\Theta_j \\ 0 \; otherwise \end{cases} \tag{10}$$

$$w_j = q_j w_j^*, \tag{11}$$

where w_j^* is the latent variable corresponding with the observed w_j; the function $f(.)$ represents the nonlinear expression to the right of the equality sign in Equation 5; q_j^* is the latent variable that corresponds with q_j, the latter derived from the censoring in budget shares; $\boldsymbol{Z}$ is a vector of control variables that could affect purchase decisions; and ν_j and ξ_j are random error terms. Assuming ν_j and ξ_j to have a bivariate normal distribution with cov $(\nu_j, \xi_j) = \delta_j$, the unconditional mean of w_j is given by (Shonkwiler and Yen 1999):[16]

$$E(w_j | \boldsymbol{X}_j, \boldsymbol{Z}_j) = \Phi(\boldsymbol{Z}_j'\Theta_j) f(.) + \delta_j \phi(\boldsymbol{Z}_j'\Theta_j), \tag{12}$$

where $\phi(.)$ and $\Phi(.)$ are the probability density and cumulative distribution functions, respectively. Denoting estimates of these functions by $\widehat{\phi(.)}$ and $\widehat{\Phi(.)}$, the stochastic food demand model corrected for censoring is given by:

$$w_j = \widehat{\Phi(.)} * f(.) + \delta_j \widehat{\phi(.)} + \zeta_j \tag{13}$$

where $\zeta_j = \xi_j - \left[\Phi(.) - \widehat{\Phi(.)}\right] * f(.) + \delta_j\left[\phi(.) - \widehat{\phi(.)}\right]$.[17] A multivariate probit via the simulated maximum likelihood estimator is used to fit the selection model in Equation 10.[18]

The third issue we address is the possible endogeneity in our proxy measure of Y in Equation 5—household food expenditure, due to the simultaneous allocation decision between total food expenditure and the budget share of the different food groups (Blundell et al. 1998) within the two-stage budgeting framework (Gorman 1959). In addition, food expenditure could be prone to measurement error due to

---

[16] The mean of w_j when $\xi_j > -\boldsymbol{Z}_j'\Theta$ is $f(.) + \delta_j \left[\frac{\phi(.)}{\Phi(.)}\right]$, whereas the mean is zero when $\xi_j \leq -\boldsymbol{Z}_j'\Theta$ (Shonkwiler and Yen 1999).

[17] The error term ζ_j has mean zero and variance $[\sigma_i^2\Phi(.) + [1 - \Phi(.)] * \{[f(.)^2] * \Phi(.) + 2f(.)\delta_j\phi(.)\} - \sigma_i^2\{\boldsymbol{Z}_j'\Theta_j * \phi(.) + \phi(.)^2\}]$, where σ_i^2 is the variance of ξ_j (Shonkwiler and Yen 1999).

[18] While single-equation probit is consistent, it will not be efficient if a household's purchase decision about a food group depends on that of another food group due to substitutability or complementarity. In this instance, estimating a multivariate probit model improves efficiency (Tauchmann 2005; Yen 2005).

errors in reported quantities or unit values used to impute own-production (Hausman et al. 1995; Lewbel 1996). This paper uses a control function approach, where a food expenditure model is first estimated using ordinary least squares (controlling for asset-based wealth index, a vector of socio-demographic variables, the other exogenous variables in the food demand model, and village fixed effects) and the estimated residual thereof is used as an additional repressor in the budget share equations as follows:

$$w_j \ = \ \widehat{\Phi(.)}\,[f(.)] + \delta_j \widehat{\phi(.)} + \rho_j \hat{e}_j^Y + \zeta_j \ \ (\forall\, j) \tag{14}$$

where $\hat{e}^Y$ is the estimated residual and Equation 14 is subject to the additional restriction $\sum_{j=1}^{n} \rho_j = 0$. Equation 14 is estimated using the iterative feasible generalized nonlinear least squares estimator, setting the initial value of α_o at 11.89 (Poi, 2012).[19] Expenditure elasticities are calculated as $\frac{\mu_j}{w_j} + 1$, while the uncompensated (Marshallian) price elasticities are calculated as $\frac{\mu_{jk}}{w_j} - \delta_{jk}$ (Banks et al. 1997) with μ_j, μ_{jk}, and δ_{jk} in turn given by:

$$\mu_j = \frac{\partial w_j}{\partial lnY} = \widehat{\Phi(.)} * \left[\beta_j + \frac{2\lambda_j}{b(\mathbf{P})}\left\{ln\left[\frac{Y}{a(\boldsymbol{P})}\right]\right\}\right] \tag{15}$$

$$\mu_{jk} = \frac{\partial w_j}{\partial p_k} \ = \ \widehat{\Phi(.)} * \left[\gamma_{jk} - \mu_j\left(\alpha_k + \textstyle\sum_{k=1}^{n} \gamma_{jk} \ln p_k\right) - \frac{\lambda_j \beta_k}{b(\mathbf{P})}\left\{ln\left[\frac{Y}{a(\boldsymbol{P})}\right]\right\}^2\right] \tag{16}$$

$$\delta_{jk} = \begin{cases} 1\; if\; j = k \\ 0\; otherwise. \end{cases} \tag{17}$$

Elasticities are computed at sample means for the total sample as well as for the beneficiary and control groups, the latter to examine possible differential responsiveness between the two groups.

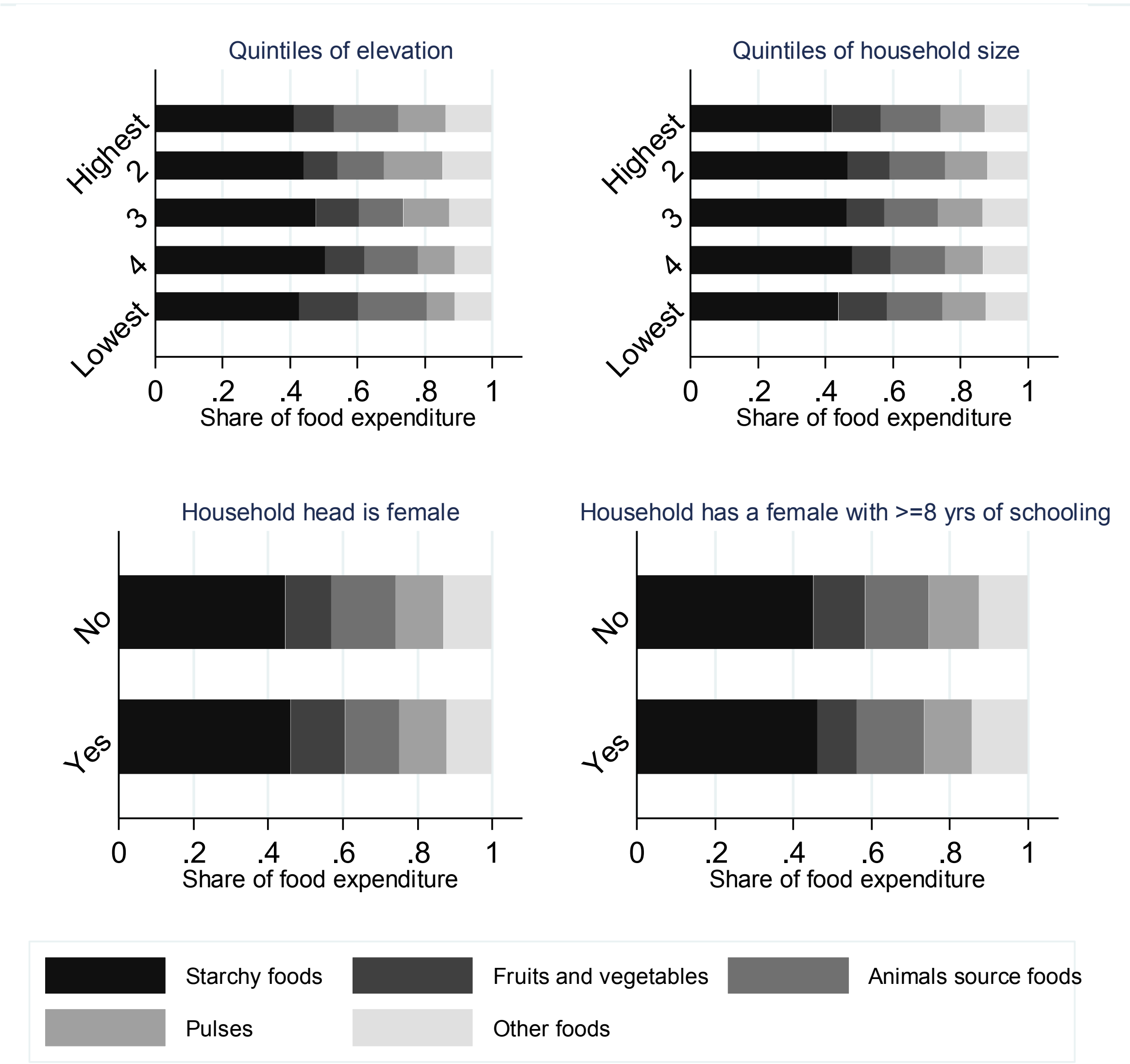
[19] The initial value of α_0 should be lower than the minimum value of the natural logarithm of food expenditure, 11.9 in our case, to ensure a positive real food expenditure (Deaton and Muellbauer 1980; Banks et al. 1997).

Table A1. Food groups

Starchy foods	**Pulses**	**Animal-source foods**	**Fruits and Vegetables**	**Other Foods**
Rice (paddy)	Peas	Goat meat		Sugar
Maize (green, cob)	Beans	Beef including minced sausage	Onions	Sugarcane
Maize (grain)	Lentils	Pork including sausages and bacon	Tomatoes	Sweets
Maize (flour)		Chicken and other poultry	Carrots	Honey, syrups, jams, marmalade, jellies, canned fruits
Millet and sorghum (grain)	Other pulses	Wild birds, insects, mice	Green pepper	Other sweets
Millet and sorghum (flour)	Groundnuts (shelled)	Other domestic/wild meat products	Other viungo	Cooking oil
Wheat, barley grain and other cereals	Groundnuts (unshelled)	Eggs	Spinach	Butter, margarine, ghee
Bread	Coconuts mature	Fresh fish and seafood	Cabbage	Other fat products
Buns and biscuits	Coconuts immature	Dried/salted fish and seafood	Other green vegetables	Salt
Macaroni, spaghetti	Bambara nuts	Package/Canned fish	Canned, dried and wild vegetables	Other spices
Other cereal products	Seeds and products from nuts/seeds (excl. cooking oil)	Other meat	Other vegetables	Tea dry
Cassava fresh	Other nuts and seeds	Fresh milk	Ripe bananas	Coffee and cocoa
Cassava (dry)		Milk products	Citrus fruits (oranges, lemon, tangerines, etc.)	Other raw materials for drinks
Cassava (flour)		Canned milk/milk powder	Mangoes	Bottled/canned soft drinks (soda, juice, water)
Sweet potatoes		Other dairy products	Avocados	Prepared tea, coffee
Yams/cocoyams			Other fruits	Bottled beer
Irish/round potatoes				Local brews
Cooking bananas, plantains				Wine and spirits
Other starches				

Source: IFPRI, 2015.

Figure A1. Expenditure shares by socioeconomic variables



Source: IFPRI, 2015.

Table A2. First-stage regression of the production models

	Treat	
	coef	se
Ownership of a nearby parcel	0.160***	0.048
Household size	0.010	0.010
Age of household head (years)	-0.000	0.001
Max adult yrs of education	-0.002	0.008
Female head	-0.038	0.034
Number of adults (age>=15)	-0.009	0.025
Distance to basic services index	-0.010	0.035
Area of parcels operated (ha)	0.104**	0.048
Owns of at least one ax	-0.054*	0.030
Non-agr. wealth index	0.012	0.019
Agricultural labor used (person-days)	0.000**	0.000
Total fertilizers applied (kg)	0.000	0.000
Travel time to seed supplier (min)	-0.000	0.001
Constant	0.156	0.130

Note: *** p<0.01, ** p<0.05, * p<0.1

District fixed effects included. Standard errors robust to arbitrary heteroskedasticity and intra-village correlation. Endogenous (instrumented) variable: participation in the program (Treat). Excluded instrument: ownership of a nearby parcel

Table A3. First-stage regression of the consumption models

	1:Treat		2:pdi		3:Treatxpdi		4:simpson_pd		5:Treatxsimpson_p		6:lnval		7:Treatxlnval	
	coef	se	coef	se	coef	se	coef	se	coef	se	coef	se	coef	se
Ownership of a nearby parcel	0.133***	0.041	0.488***	0.144	0.786***	0.251	0.040**	0.017	0.041**	0.018	0.067	0.092	1.575***	0.471
Ln of altitude (meters)	-0.279*	0.166	0.375	0.364	-1.979**	1.004	0.180***	0.037	-0.097	0.077	0.661**	0.283	-3.251*	1.905
Number of intercropped plots	0.064***	0.013	0.652***	0.060	0.709***	0.104	0.020***	0.006	0.033***	0.006	0.140***	0.024	0.835***	0.161
Received extension service (last 12 months)	0.454***	0.054	0.875***	0.136	2.754***	0.329	0.034**	0.013	0.183***	0.024	0.319***	0.119	5.274***	0.613
Household size	-0.007	0.016	0.068	0.058	0.028	0.099	-0.013**	0.007	-0.005	0.008	0.084**	0.038	-0.063	0.185
Age of household head (years)	0.000	0.001	0.003	0.004	0.002	0.007	0.000	0.000	-0.000	0.000	-0.001	0.003	0.003	0.013
Max adult yrs of education	-0.000	0.006	-0.013	0.026	-0.027	0.042	0.002	0.002	0.001	0.003	0.038**	0.018	0.021	0.069
Female head	-0.021	0.034	-0.072	0.131	-0.075	0.200	-0.036**	0.015	-0.022	0.017	0.052	0.084	-0.253	0.394
Number of children (age<15)	0.017	0.019	0.083	0.069	0.097	0.122	0.020***	0.007	0.010	0.009	-0.041	0.050	0.204	0.224
Has a female with at least eight yrs of schooling	-0.079*	0.045	0.048	0.203	-0.320	0.335	-0.008	0.016	-0.034	0.022	-0.202**	0.091	-1.058*	0.540
Distance to basic services index	0.010	0.025	0.031	0.070	0.083	0.157	0.007	0.007	0.006	0.010	-0.063	0.038	0.108	0.290
Area of parcels operated (ha)	0.077*	0.042	0.669***	0.103	0.536**	0.270	0.051***	0.011	0.041**	0.020	0.358***	0.088	0.931*	0.488
Owns of at least one ax	-0.054*	0.028	0.461***	0.117	0.055	0.190	0.028**	0.012	-0.005	0.013	0.181***	0.059	-0.539	0.331
Non-agr. wealth index	0.022	0.017	0.196**	0.078	0.219	0.134	0.003	0.007	0.005	0.009	0.091**	0.037	0.301	0.213
Experienced food shortage (June-Sep 2013)	-0.036	0.044	-0.607***	0.138	-0.265	0.219	-0.050***	0.016	-0.013	0.019	-0.433***	0.110	-0.550	0.482
_cons	1.768	1.101	-0.955	2.467	11.650*	6.628	-0.970***	0.263	0.607	0.507	5.418***	1.987	19.975	12.589
Number of observations	935		935		935		935		935		935		935	

Note: *** p<0.01, ** p<0.05, * p<0.1

Treat, pdi, Treatxpdi: endogenous (instrumented) variables in the food consumption model when dependent variable is ddi.

Treat, simpson_pd, Treatxsimpson_pd: endogenous (instrumented) variables in the food consumption model when dependent variable is simpson_dd.

Treat, lnval, Treatxlnval: endogenous (instrumented) variables in the food consumption model when dependent variable is lnfexp.

Excluded instruments: ownership of a nearby parcel, altitude, number of intercropped plots, and received extension service (last 12 months).