The Effects of Smallholder Agricultural Involvement on Household Food Consumption and Dietary Diversity: Evidence from Malawi

Rui M. S. Benfica
Strategy and Knowledge Department
International Fund for Agricultural Development (IFAD), Rome, Italy

Talip Kilic
Development Research Group
World Bank, Washington, DC, USA

Abstract. *We investigate how household agricultural involvement affects food consumption and dietary diversity in Malawi. Ceteris paribus, a 10% increase in on-farm income share increases food consumption/capita by 2.9%, calorie intake/person/day by 1.7%, and leads to small improvements in dietary diversity. There are significant differences in the relationship between on-farm income shares and caloric shares: a positive and significant relationship with the shares from energy dense/low protein cereals/grains, but not significant with shares from nuts/pulses and sugars. Negative relationships are found with shares from roots/tubers, vegetables/fruits, oils/fats, and meat/fish/milk. While food consumption and dietary diversity increase with agricultural involvement, the quality of diets is an issue. As purchased calories are associated with richer/high quality diets, particularly protein rich, households with lower dependency on agriculture meet those diets more easily, highlighting the importance of crop and income diversification to dietary diversity. Nutrition education and crop diversification programs can improve food security and nutritional outcomes.*

Keywords: Agricultural involvement, consumption, dietary diversity, and endogeneity.

JEL codes: C31, C36, D12, D13, D24, and I13.

The Effects of Smallholder Agricultural Involvement on Household Food Consumption and Dietary Diversity: Evidence from Malawi

1. Introduction

Malawi is a predominantly rural economy with agriculture accounting for 30 percent of the Gross Domestic Product, 84 percent of households owning and/or cultivating land, and the overwhelming majority of farming households practicing subsistence agriculture. In spite of significant public spending in agricultural development programs in recent years, rural poverty and food security remain high with over half of the population living below the poverty line. Over the period 2005 to 2011, income growth has been significantly regressive, the poverty headcount remained stagnant, and income inequality has risen (World Bank, 2014).

In recent years, as a result of publicly funded agriculture support programs, there have been significant increases in the levels of intensification, with more households using productivity enhancing inputs. Indications are, however, that that has been accompanied by increased crop specialization, especially towards maize grain. In fact, except for maize, pigeon peas and tobacco, the proportion of households growing all other crops has fallen between 2005 and 2011. Research indicates that, overall, income diversification away from own-farm has also fallen. Except for paid farm work (a relatively low return activity), off-farm income diversification fell significantly over the period (World Bank, 2014).

As a result of this dynamic, the importance of crop income in total household income has increased in recent years. The growth in crop income has, however, less than compensated for the loss in income from non-farm sources, which implied stagnant consumption poverty reduction in rural areas. Those patterns can have important implications for food security and nutrition. Depending on how households procure their calories from different sources, between own production and market purchases, there can be implications for overall levels and, more importantly, to dietary diversity patterns.

In terms of nutritional outcomes, while there have been improvements in recent years, children and household level nutrition insecurity remains high. By 2011, 23% of Malawian households had inadequate food consumption (poor and borderline, by WFP standards), with female-headed and poorer rural households exhibiting higher incidence of food inadequacy. In terms of child nutritional outcomes, there is high prevalence of stunting rates (29.9%), modest underweight rates (6.9%), but relatively low wasting rates (3.8%).

In the context of the increased importance of agriculture relative to non-farm income, the persistent and high levels of poverty, and food and nutrition insecurity, the links between agriculture and nutrition are potentially important. However, those links have not received adequate attention to date mainly due to data limitations. This paper uses data from the Malawi Integrated Household Survey 2010/11 (IHS3) to start filling that gap by investigating the effect of agricultural involvement - defined as the share of on-farm income in total income - on household consumption and dietary diversity.

2. Analytical Framework and Research Questions

The term "nutrition-agriculture linkages" refers to a set of relationships that describe the mutual dependence of nutrition, health and agriculture. The Nutrition-agriculture framework features looping relationships that illustrate the bi-directional causality, and thus interdependence, among their key components (Chung 2012). Changes in nutrition or health status are expected to affect agricultural production; conversely changes in the agricultural sector can have significant effects on individual health and nutritional status (Sahn 2010). Figure 1 summarizes a basic framework for analyzing agriculture nutrition linkages. It highlights the interdependent relationship by focusing on rural households and highlighting the relationships that connect nutrition, agriculture and health at the household and individual levels.

Given the question in hand, this paper takes a more simplified approach. Our focus is on nutritional outcomes, and as a result, we do not detail the ways in which agriculture affects health status, and indirectly, nutritional status. Hoddinott (2011) describes in detail that more complete loop. Here, we start from the trickle down approach that assumes that an increase in output will elicit changes in household nutritional status (Chung 2012). Nutritional status is presumed to improve as a result of increases in own consumption or income. The trickle-down strategy can also benefit net consumers if aggregate production changes are large enough to reduce the price of crops that are nutritionally important.

The left-hand side of Figure 1 shows that household food production is expected to improve individual food intake by either (a) increasing consumption from own-production or (b) contributing to household income for the purchase of food. In turn, improved food intake provides energy that is needed for bodily growth, maintenance and activity. A high quality diet also provides protein and various micronutrients (vitamins and minerals) that are essential for optimum growth and functioning (Chung, 2012; TFCSD, 1991).

Since agricultural activity determines, to a great extent, the amount, type, stability, control, and distribution of income, the linkages among agriculture and consumption are expected to be strong and direct for agricultural households (Chung, 2012). Furthermore, agriculture affects the food available for consumption by the household, including its diversity, quality and price (von Braun et al 2010; Chung, 2012).

In this paper, we argue that weather increased output or the importance of agricultural income will result in increased consumption and improved nutritional outcomes, is an empirical question that needs to be tested in each context and particular circumstance. So, we ask and seek to answer the question: What are the effects of rural agricultural involvement of rural Malawian households on consumption, calorie intake and dietary diversity?

3. Data

The analysis uses household level data from the Malawi Integrated Household Survey III (IHS3). The survey was conducted by the National Statistical Office (NSO), supported by the Living Standards Measurement Surveys – Integrated Surveys in Agriculture (LSMS-ISA) project at the World Bank, from March 2010 through March 2011. The sample includes 12,271 households (10,038 from rural and 2,233 from urban areas). The sampling design is representative at the national, rural and urban, and district level hence the survey provides reliable estimates for those areas. It covers topics ranging from household demographics, consumption patterns and expenditure levels, agricultural, livestock, and fisheries production and marketing, child anthropometry, among other variables (NSO, 2011).

For the purposes of this analysis we rely on the rural agricultural household sample of around 9,000 households, corresponding to approximately 92 percent of the overall rural sample. It uses survey data to generate variables related to the level of agricultural involvement (reliance on agricultural income) and food consumption and nutritional outcomes at the household level.

4. Defining Agricultural Involvement, and Food Consumption and Nutritional Outcomes

Agricultural involvement is defined as household on-farm (crop and livestock) income as a share of total gross household income. This measure captures the relative weight of returns to household agricultural involvement. The higher the share of on-farm income, the lower the level of household income diversification, i.e., the share of income generated from off farm sources. Data from IHS2 and IHS3 indicate that between 2005 and 2011, household agricultural involvement has increased, remaining at relatively high levels, with both the share of households engaging in on-farm activities and the share of net income from those sources increasing significantly (Table 1).[1]

The range of outcome variables that inform our analysis include (a) household annual food consumption expenditures per capita, (b) household caloric intake per capita per day, (c) household food consumption score, (d) household Simpson Index of dietary diversity, and (e) shares of caloric intake attributed to (i) cereals and grains, (ii) roots and tubers, (iii) nuts and pulses, (iv) vegetables and fruits, (v) meat, fish, milk, and other animal products, (vi) oils and fats, (vii) sugar products, and (viii) other food items. The following is a basic definition of each of these outcomes.

Food consumption is measured both in terms of total value of food consumed and the corresponding calories levels consumed per household.

Food consumption per capita ($FCpc_h$) is defined as the total value of food consumed in the household annually divided by household size. It can be represented as:

$$(1) \qquad FCpc_h \approx \frac{1}{N_h}\sum_{i=1}^{n} C_{hi}$$

[1] The rural sample here includes all rural households. The analysis in this paper only considers rural agricultural households. The results in the bottom panel of Table 1 refer to the shares of net income, while the analysis in the paper focuses on the share of on-farm income in gross total household income.

where, C_{hi} is value of household *h* annual consumption of commodity *i*, N_h if the size of household *h*. This measure is expressed in monetary terms.

Calorie intake per capita per day ($Calpcpd_h$) is computed by converting the annual quantities of individual food items consumed to calories using standard conversion factors. The sum of calories across all food items is then divided by household size and 365 days to get to the daily level of per capita calorie consumption.

$$\text{(2)} \qquad Calpcpd_h \approx \frac{1}{365 * N_h} \sum_{i=1}^{n} Cal_{hi}$$

where, Cal_{hi} is calorie consumption of food item *i* by household *h*, N_h if the size of household *h*. This measure is expressed as number of calories.

Dietary diversity is measured using the Food Consumption Score, the Simpson Diversity Index, and the *share of calories from food groups*. The definitions are as follows.

The Food Consumption Score (FCS_h) is a composite score based on dietary diversity, food frequency, and relative nutritional importance of different food groups. Food items consumed in the 7 days prior to the interview are grouped into 8 groups. The consumption frequency (maximum of 7 days/week) of each food group by the household is then multiplied by group assigned nutrient based weights. Those values are then summed up to make the household *h* FCS.

$$\text{(3)} \qquad FCS_h \approx \sum_{i=1}^{8} f_{hi} * w_i$$

where f_{hi} is the frequency of consumption of food commodity group *i* by household *h*, and

and w_i is the weight attributed to each food commodity group *i*. The World Food Program (WFP) proposed this indicator.[2]

[2] The score thresholds range from 0 – 35 and allows for the classification of households into the following categories of food consumption: (1) Poor (FCS between 0 -21), (2) Borderline (FCS between 21 and 35), and Acceptable food consumption (FCS above 35). The WFP calls inadequate consumption to the combination of poor and borderline i.e., FCS less than 35.

The Simpson diversity index (SDI_h) is a member of a class of diversity indexes that take into account, not only whether or not each food item is consumed, but also the relative importance of each type of food consumed, as expressed by consumption shares. It can be expressed as

$$(4) \qquad SDI_h = 1 - \sum_{i=1}^{n} ShCal_{hi}^2$$

where, $ShCal_{hi}$ is the calorie consumption share of food item *i* in total calorie consumption of household *h*, and *n* is the total number of food items considered. By considering the shares of calorie consumption, it implicitly gives more weight to food types that have higher shares. Food items with equal shares are weighted equally.[3] This index ranges from 0 to 1. If a household consumes only one type of food item, i.e., its share is equal to unity, the index will be zero (no diversity). As more items are consumed the index value will increase indicating more dietary diversity.

The food groups defined in this analysis are formed in line with the structure suggested by the WFP for the FCS. Annex Table A1 lists the crops and products comprising the groups, and a description of their nutritional attributes.

The analysis uses shares of calorie consumption from these groups as outcomes in the 3SLS model to access how agricultural involvement affects the relative levels of those shares. The share of calorie consumption from each food group *i* ($ShCal_{hi}$) can be expressed as

$$(5) \qquad ShCal_{hi} \approx \frac{Cal_{hi}}{\sum_{i=1}^{n} Cal_{hi}}$$

where, Cal_{hi} is the calorie consumption share of food item *i* of household *h*, and *n* is the total number of food items considered to arrive at the total number of calories consumed (denominator). By definition the sum of the shares will be equal to unity.

[3] In a more elaborate scheme, one may want to give bigger weights to items such as vegetables, meat and fish (and very small weights to items such as sodas, cookies and alcohol) as opposed to staple foods.

5. Descriptive Statistics of Agricultural Involvement and Food Consumption and Nutritional Outcomes

This section looks at some descriptive statistics for the variables of interest. Using IHS3 survey data, we look at (a) levels of agricultural involvement and household level consumption, calorie intake and dietary diversity outcomes by selected household characteristics, such as gender of the head and poverty status, and rural-region of residence. Results are presented in Table 2.

Overall, in rural Malawi, where about 92 percent of the households engage in crop and livestock production, the share of on-farm income from total gross household income is about 60%, i.e., for every Malawi Kwacha generated about 60 cents originate from that source. The following results stand out. First, agricultural involvement, defined by this measure, is higher (at 66%) in the central region and lowest (at 53%) in the south. Second, while differences are relatively small, female headed and poor households (63%) have relatively higher levels of agricultural involvement than their male (59%) and non-poor (57%) counterparts. On average, households in the top 20% of income have shares of on-farm income of only about 50%, compared to about 61% among the poorest 20%.

Descriptive analyses of the consumption and dietary diversity outcomes indicate several important results. First, food consumption per capita is higher in the central region and lowest in the south, pretty much in line with the patterns of agricultural involvement. Second, contrary to that, when looking at consumption per capita across gender and wealth, we find that male headed and non-poor households (that exhibited lower levels of agricultural involvement) enjoy relatively higher levels of consumption per capita. Third, calorie intake and dietary diversity, measured through the food consumption score and the Simpson Index, are higher in the north, among male-headed, and relatively wealthier households. One exception is calorie consumption per capita per day that is slightly higher among female-headed households when compared to their male counterparts (Table 2). Annex Table 2 provides a more detailed analysis of food consumption adequacy (for all rural households, by gender of headship, and poverty status) derived from the food consumption score. By 2011, 23% of Malawian households had inadequate food consumption (poor and borderline, by WFP standards), with female-headed and poorer rural households exhibiting higher incidence of food consumption inadequacy.

Finally, when looking at disaggregated calorie consumption, we note that households in the center and south have a structure essentially dominated by cereals, 72% and 68%, respectively, significantly small shares from non-crop protein sources such as meat/fish/milk (just about 3%) and less than 6% from roots/tubers, while household in the north appear to have a relatively more balanced diet, deriving about 13% of calories from meat/fish/milk, and just below 60% from cereals/grains, and 12% from roots/tubers. While differences are not significant, male-headed households and those that are classified as non-poor enjoy relatively more balanced diets, consuming relatively less calories from non-cereal sources. A look at the structure of calorie consumption by wealth quintile reveals that the poorest 20% derive about 80% of their calorie consumption from cereals, against only 61% among the top 20% richest, that have in turn a relatively more diversified diet where the share of fruits and vegetables and meat/fish/milk is about double of that of the poorest 20%.

In terms of food consumption and diversity outcomes across different levels of agricultural involvement we find that in a strictly bivariate sense, i.e., without controlling for a wealth of factors: (a) higher levels of agricultural involvement almost invariably result in lower levels of the various aggregate outcomes. While this seems to represent an apparent paradox, an analysis on a more detailed continuum shows that beyond a certain involvement threshold a positive relationship holds beyond shares of over 50%, levels at which over 60% of the sample falls (Figure 2). For consumption per capita, calories per capita per day and food consumption score a clear mirrored J-shaped curve emerges; (b) when looking at the different calorie sources, we find that the share of cereals/grains (mostly sourced from own production) increases with the levels of agricultural involvement, while the shares of calories from most of the other sources fall, particularly for those that are mostly sourced from the market. This reflects the degree of difficulty households that are increasingly specialized in agriculture have to acquire calories from market sources, especially non-crop protein sources such as meat/fish/milk and oils/fats (Table 3 and Figure 3).

Assessing the true average effect of agricultural involvement can only be done by controlling for household and location factors, while addressing the potential endogeneity of agricultural involvement. The section that follows accomplishes that.

6. Econometric Methods

Given the presence of unobserved heterogeneity that may jointly determine the dependent variables and the explanatory variable of interest, we rely on Two-Stage Least Squares (**2SLS**) regressions for the analysis of the effects of agricultural involvement on consumption and nutritional outcomes, and a simultaneous system of equations in a Three-Stage Least Squares (**3SLS**) framework for the analysis of the effects of agricultural involvement on caloric shares from the different food groups. The regressions control for a rich set of household and community characteristics, combined with geospatial variables that broadly capture climatological conditions, soil characteristics, and agricultural productivity potential obtained by linking geo-referenced household locations to publically available geographical information systems. We present each model at a turn.

6.1. Two-Stage Least Squares Model (2SLS)

This model is used to estimate the effects of agricultural involvement on household level food consumption levels, and nutritional and dietary diversity outcomes, while controlling for a wealth of household and district level characteristics. Equations (3) and (4) represent the 2SLS model for each Outcome Variable ($Y_h^{\,j}$).

(6) $Y_h^{\,j} = \alpha_0 + \alpha_1 X_{1h} + \alpha_2 A_h + \varepsilon_h$

(7) $A_h = \beta_0 + \beta_1 X_{1h} + \beta_2 X_{2h} + \eta_h$

> *Outcome indicators* ∈ *j* ={food consumption per capita; calorie consumption per person per day; food consumption score; the Simpson Diversity Index}

Where $Y_h^{\,j}$ is household *h* food consumption or nutritional outcome *j* as described in the previous section (expressed in logarithm form), A_h is the endogenous variable – share of on-farm income representing agricultural involvement of household *h* (expressed in logarithm form), X_{1h} is a vector of exogenous variables assumed to be associated with consumption or nutritional outcomes and agricultural involvement. They include household characteristics, such as head's gender, age and education, household size, income diversification, access/use of services; farm characteristics, and location specific fixed-effects factors. X_{2h} is a vector of instrumental variables for agricultural involvement. ε_h and η_h are error terms, $E(\varepsilon_h) = 0$, $E(\eta_h) = 0$, and cov $(\varepsilon_h; \eta_h) = 0$. The analysis runs the model for each individual consumption and nutritional outcome *j*, separately.

The instrumental variable (IV) used to capture the random variation in the share of household on-farm income that is not directly related to the dependent variables was number of agricultural officer per household at the district level. A key requirement for the validity of the instruments is that they are sufficiently strongly correlated with the endogenous variable and uncorrelated with the error term. In other words, they do not affect consumption or nutritional outcomes directly, but only through agricultural involvement.

6.2. Three-Stage Least Squares (Simultaneous Equations) Model (3SLS)

The 3SLS simultaneous equations system of the share of on-farm and the food group shares of calorie intake is used to look at the effects of agricultural involvement on decomposed calorie intake by looking at individual food groups (grains, roots, pulses, fruits and vegetables, oils/fats and sugars).[4] The 3SLS model can be represented as

(8) $ShCal_{hi} = \alpha_{0i} + \alpha_{1i} X_{1h} + \alpha_{2i} A_h + \varepsilon_{hi}$, for each food group *i*

(9) $A_h = \beta_0 + \beta_1 X_{1h} + \beta_2 X_{2h} + \eta_h$

(10) $$\sum_{i=1}^{n} ShCal_{hi} = 1$$

Food groups ∈ A*i* ={Cereals, Roots, fruits/vegetables, meats/fish/milk, oils/fats, sugars, and other}

Where $ShCal_{hi}$ is household *h* share of calories from food group *i (i equations)*, A_h is the endogenous variable (share of on-farm income representing agricultural involvement of household h), X_{1h} is a vector of exogenous variables, and X_{2h} is a vector of instrumental variables for agricultural involvement. ε_h and η_h are error terms as defined earlier. The equations

[4] As described in Table 2, these groups are differentiated by relative energy density, protein content, and absorbability of micronutrient content.

are estimated as a system of *j* share equations, an agricultural involvement equation (share of on-farm income, satisfying constraint (10). The estimate of interest is α_2 for each *i.*

7. Results

As discussed in section 3, we use 2SLS and 3SLS techniques to address the endogeneity of our variable of interest, i.e., unobserved heterogeneity that may jointly determine consumption and dietary diversity outcomes and agricultural involvement. The 2SLS model is used to analyze the effects of agricultural involvement on the levels of consumption and dietary diversity outcomes, while the 3SLS system of equations analyses the effects of agricultural involvement on the shares of calories consumed from the various food groups.

7.1. Choice of Instruments

Several potential instruments were considered. A key question we tried to answer to come up with appropriate instruments was whether it was reasonable to consider a set of instruments at the district level, i.e., if the unobserved heterogeneity was not at the district level, but rather at lower levels (enumeration area, or household). To test that, we run regressions of selected outcomes on on-farm income share, a wide range of household level factors and district dummies. Essentially, we run two models, a model with district fixed-effects (FE) and another with district random effects (RE) followed by a Hausman test to conclude if the difference in coefficients was systematic (FE) or not (RE).

The test results (Table 4) indicate that we reject district fixed-effects, i.e., there is not unobserved, time invariant heterogeneity at district-level. So, unobserved heterogeneity is more likely at finer levels such as EA or household. District level Instrumental Variables (IVs) are, therefore, appropriate to be used in our models.

The final models chosen are *just identified*. The instrumental variable (IV) used to capture the random variation in the share of household on-farm income that is not directly related to the dependent variables is district-level of agricultural extension officers per household. These data are obtained from the IHS3 survey, and records from the Ministry of Agriculture and Food

Security, respectively.[5] To verify the appropriateness of our models in addressing endogeneity through IV (i.e., adequacy of the instruments), in addition to checking the magnitude and the statistical significance of the correlation with the endogenous variable, we run post-estimation diagnostic tests, such as Wu-Hausman endogeneity, and Cragg-Donald for weak-identification.

7.2. 2SLS Model Results

The 2SLS Model indicates several results (Table 5). First, first stage regression results indicate that there is a strong positive and statistically significant correlation between the share of on-farm income and the chosen instrument (district-level of agricultural extension officers per household), which is a necessary requirement for its adequacy.

Second, there are several other factors strongly associated with the share of on-farm income. Households have lower levels of agricultural involvement if they are male headed, achieved relatively high levels of education, have more diversified sources of income off the farm, and are relatively wealthier overall. Higher levels of agricultural involvement are associated with number of female adults, use of agricultural extension and input use, high levels of agricultural asset ownership (agricultural asset index), and prevalence of severe nutrient availability constraints.

Finally, more importantly, the results from the second stage **2SLS** estimations indicate that, controlling for head characteristics, household composition, agricultural technology, income diversification, and region-specific fixed-effects, on average, a 10% increase in the share of on-farm income: (a) increases food consumption per capita by 2.9 percent (900 Kwachas per capita) and total calorie intake per capita per day in 1.7 percent (41 calories per person per day); and (b) leads to small improvements in dietary diversity, resulting in an increase of 1.02 percent in the food consumption score and 0.97 percent in the Simpson Index. These positive effects are statistically significant at least at the 5 percent level.

In each of the **2SLS** estimations, we reject the exogeneity of the on-farm income share (the main explanatory variable), justifying therefore the use of the IV approach. In each case, we also find evidence that counteracts potential concerns regarding weak instrumental variable bias (Table 5).

[5] We decided for a just identified model after attempting the use of multiple district level instruments (e.g., covariance of average precipitation in rainy season, inputs distributed per household, number of lead farmers per household, among others). Tests of over-identifying restrictions (Sargan Chi2) have systematically not supported the validity of additional instruments.

7.3. 3SLS Simultaneous Equation Model Results

The 3SLS simultaneous equations model is aimed at assessing the extent to which agricultural involvement relates to the structure of household consumption. More specifically, we evaluate how increased share of on-farm income affects the shares of calories consumed from the defined food groups (grains, roots, pulses, fruits and vegetables, oils/fats and sugars) whose nutrition attributes are described in Annex Table 1.

The **3SLS** estimations (Table 6) reveal fundamental differences in the relationship between the caloric shares and share of household on-farm income. First, there is a positive and statistically significant relationship between the share of household net on-farm income and the shares of calories from energy dense and low protein grains and cereals. A positive effect is also observed on sugars, but it is not statistically significant. Second, there is a negligible and not statistically significant impact on the share of caloric intake associated with nuts and pulses. Finally, we recover negative and statistically significant relationships between the share of household net on-farm income and the shares of calories from (a) roots and tubers, (b) vegetables and fruits, (c) oils and fat, and (d) meat, fish and milk products (high quality protein/easily absolvable micronutrient foods).

8. Conclusions and Policy Implications

In the context of persistent and high levels of poverty and food insecurity in Malawi, the links between agriculture and nutrition are potentially important, but have not received adequate attention. In part, that has been due to data limitations. Taking advantage of data recently collected in Malawi (IHS3) allowing for linking household consumption with the structure of economic activity and income sources, this paper starts filling that gap by investigating the effect of agricultural involvement on household consumption and dietary diversity.

The analysis uses a rural agricultural household sample of over 9,000 households, i.e., 92 percent of the overall rural sample. The range of outcome variables that inform our analysis include (a) household annual food consumption expenditures per capita, (b) household caloric intake per capita per day, (c) household food consumption score, (d) household Simpson index of dietary diversity, and (e) shares of caloric intake attributed to (i) cereals and grains, (ii) roots and tubers, (iii) nuts and pulses, (iv) vegetables and fruits, (v) meat, fish, milk, and other animal products,

(vi) oils, (vii) sugar products, and (viii) other food items. The main explanatory variable is net household income from crop and livestock activities as a share of total net household income, a measure that captures the relative weight of returns to household agricultural involvement.

Given the presence of unobserved heterogeneity that may jointly determine the dependent variables and the explanatory variable of interest, we rely on Two-Stage Least Squares (**2SLS**) regressions for the analysis of outcomes (*a)* through (*d)*, and a simultaneous system of equations in a Three-Stage Least Squares (**3SLS**) framework for the analysis of caloric shares, as defined above. The regressions control for a rich set of household and community characteristics, combined with geospatial variables that broadly capture climatological conditions, soil characteristics, and agricultural productivity potential and that are obtained by linking geo-referenced household locations to publically available geographical information systems.

Conditional on observable attributes that are part of our models, we provide results from diagnostic tests that reject the presence of unobserved district-level heterogeneity, in support of district level IVs. The models are just-identified, and the instrumental variable (IV) used is district-level of agricultural extension officers per household. This variable is strongly correlated with the share of on-farm income. Through post estimation tests, we reject the exogeneity of the main explanatory variable (justifying the need for the 2SLS approach), and provide evidence that counteracts potential concerns regarding weak instrumental variable bias.

Controlling for a wealth of household and region-specific factors, **2SLS** estimations indicate that, on average, a 10% increase in the share of on-farm income: (a) increases food consumption per capita by 2.9 percent and total calorie intake per capita per day in 1.7 percent; and (b) leads to only small improvements in dietary diversity, resulting in an increase of 1.02 percent in the food consumption score and 0.97 percent in the Simpson Index. These positive effects are statistically significant at least at the 5 percent level.

The **3SLS** estimations reveal fundamental differences in the relationship between the caloric shares and share of household net on-farm income. While there is a positive and statistically significant relationship between the share of household on-farm income and the shares of calories from energy dense/low protein cereals/grains, there is no statistically significant impact on the caloric intake associated with nuts and pulses and sugars. Furthermore, we find negative and statistically significant relationships between the share of household on-farm income and the shares of calories consumed from (a) roots and tubers, (b) vegetables and fruits, (c) oils and fat, and (d) meat, fish and milk products (high quality protein/easily absolvable micronutrient foods), that normally define more diversified diets that are predominantly purchased.

These results indicate that although household food consumption and dietary diversity increases with agricultural involvement, there are issues related to the quality of the diets, as energy dense diets increase with agricultural involvement. As purchased calories are associated with richer and high quality diets (particularly protein rich diets), households with lower degrees of dependence on agriculture seem to be able to meet those diets more easily, highlighting the importance of overall income diversification to rural livelihoods. It should highlighted, however, that programs aimed at improving household nutrition education and promoting crop diversification at the farm level should also be considered as a way to ensure that households achieve poverty reduction, while improving household food and nutrition security.

References

Chung, K. 2012. Introduction to Agriculture Nutrition Linkages. Working Paper No. 72E. Directorate of Economics. Ministry of Agriculture, Maputo, Mozambique.

Hoddinott, J. 2011. Agriculture, Health, and Nutrition. Toward Conceptualizing the Linkages. Leveraging Agriculture for Improving Nutrition and Health, 2020 Conference Paper 2.

National Statistics Office (NSO-Malawi). 2011. "Integrated Household Survey 2010-2011." Zomba, Malawi.

Sahn, D. 2010. The Impact of Poor Health and Nutrition on Labor Productivity, Poverty, and Economic Growth in Sub-Sahara Africa. In P. Pinstrup-Andersen, ed., The African Food System and Its Interaction with Human Health and Nutrition. Ithaca, NY: Cornell University Press in cooperation with the United Nations University.

Task Force for Child Survival and Development. 1991. Proceedings of Ending Hidden Hunger (A Policy Conference on Micronutrient Malnutrition). Montreal, Quebec.

Von Braun, J, M. Ruel and S. Gillespie. 2010. Bridging the Gap: Linking Agriculture, and Health to Achieve the Millennium Development Goals. In P. Pinstrup-Andersen, ed., The African Food System and Its Interaction with Human Health and Nutrition. Ithaca, NY: Cornell University Press in cooperation with the United Nations University.

World Bank. 2014. Trends and Determinants of Household Welfare in Malawi: Implications for Policy. Washington DC.

World Food Programme. 2007. Food consumption analysis: Calculation and use of the Food Consumption Score in food consumption and food security analysis (draft). Rome.

Tables and Figures

Tables

Table 1. Trends in income diversification, 2005-2011

Selected Indicators	Rural Areas		
	2005	2011	Difference
Income Sources (% of HHs)			
Agricultural			
Crop and livestock production	83.4	92.2	+8.8**
Agricultural wage	54.4	48.7	-5.7**
Farm rents	2.2	0.50	-1.7**
Non-Agricultural			
Self-employment	29.8	16.5	-13.3**
Non-farm wage	16.2	13.2	-3.0**
Non-farm rents	2.3	1.9	-0.4
Net Income Shares (% of Income)			
Agricultural			
Crop and livestock production	65.4	71.3	+5.9**
Agricultural wage	11.3	15.8	+4.5**
Farm rents	0.1	0.0	-0.1**
Non-Agricultural			
Self-employment	8.8	5.0	-3.8**
Non-farm wage	7.5	7.6	+0.1**
Non-farm rents	0.1	0.4	+0.3**

Note: Significance level of the difference: 1% (**), 5% (*), and 10% (+).
Source: World Bank, 2014 using Malawi IHS2 and IHS3.

Table 2. Agricultural Involvement and Household Level Outcomes, by selected characteristics

Household characteristics	Agricultural Involvement (Share of On-Farm Income)	Household Level Consumption and Dietary Diversity Outcomes (Rural Agricultural Households)											
		Food and Calorie Consumption		Dietary Diversity		Disaggregated Calorie Consumption: Share of Calorie Consumption per person per day by Food Groups (%)							
		Food Cons. per capita	Calories/ person/day	Food Cons. Score	Simpson Index	Cereals/ Grains	Roots/ Tubers	Pulses/Nuts	Fruits/ Vegetables	Meat/Fish and Milk	Oils/ Fats	Sugars	Others
Rural Malawi	0.60	31,044	2,425	49.6	0.56	68.5	6.5	5.3	2.5	4.6	5.7	5.4	1.6
By Region													
North	0.64	31,041	2,595	51.2	0.57	56.4	12.8	3.7	1.7	14.0	5.0	5.2	1.2
Center	0.66	32,748	2,361	49.5	0.55	72.4	5.0	5.4	2.1	3.1	5.1	5.1	1.8
South	0.53	29,560	2,433	49.3	0.57	68.4	6.2	5.5	3.1	3.3	6.4	5.6	1.6
By Sex													
Male	0.59	31,369	2,396	50.0	0.58	67.7	6.4	5.3	2.4	4.9	6.0	5.5	1.7
Female	0.63	30,053	2,513	45.4	0.52	70.7	6.9	5.1	2.7	3.7	4.7	4.9	1.3
Consumption Quintiles													
1st Quintile	0.61	11,404	1,360	36.9	0.39	78.4	5.3	3.3	2.4	2.9	4.2	2.9	0.6
2nd	0.64	18,346	1,906	43.3	0.50	72.0	6.5	4.5	2.1	3.8	5.7	4.3	1.1
3rd	0.64	25,392	2,299	49.0	0.58	67.9	6.5	5.4	2.6	4.2	6.0	5.7	1.7
4th	0.60	35,423	2,760	55.3	0.64	63.9	7.1	6.2	2.7	5.2	6.4	6.5	2.0
5th Quintile	0.50	64,885	3,814	63.8	0.71	60.0	7.4	7.0	2.7	6.9	6.0	7.4	2.7
Poverty Status													
Poor	0.63	16,458	1,734	41.4	0.46	74.5	6.0	4.1	2.3	3.5	5.2	3.9	1.0
Non-Poor	0.57	45,329	3,001	57.7	0.66	62.9	7.1	6.4	2.7	5.7	6.1	6.8	2.2

Source: Malawi IHS3 Survey.

Table 3. Household Level Outcomes by Agricultural Involvement

Household Level Consumption and Dietary Diversity Outcomes	Levels of Agricultural Involvement (Terciles of Share of On-Farm Income)			All Households
	Low	Middle	High	
	0.15	0.64	0.98	0.60
Consumption and Diversity Measures				
Food and Calorie Consumption				
Food Consumption per capita	35,077	29,080	29,037	31,044
Total Calories per person per day	2,513	2,402	2,362	2,425
Dietary Diversity Measures				
Food Consumption Score	52.7	48.1	48.2	49.6
Simpson Diversity Index	0.58	0.57	0.54	0.56
Share of Calorie Consumption by food groups (%)				
Grains	67.0	68.4	69.7	68.5
Roots	6.4	6.9	6.4	6.5
Pulses	5.4	5.1	5.2	5.3
Fruits/Vegetables	2.5	2.7	2.4	2.5
Meat/Fish/Milk	4.9	4.3	4.5	4.6
Oils/Fats	6.5	5.5	5.1	5.7
Sugars	5.6	5.3	5.2	5.4
Others	1.8	1.8	1.3	1.6

Source: Malawi IHS3 Survey.

Table 4. Testing District Random vs. Fixed Effects

Outcome Variables	Ho: Difference in coefficients not systematic (RE)		
	Chi2 (1)	Prob>chi2	Unobserved, time-invariant heterogeneity at district-level?
Household Level			
Food Consumption per Capita	8.78	0.845	No
Calories pc per day	3.63	0.993	No
Food Consumption Score	6.13	0.963	No

Source: Malawi IHS3 Survey.

Table 5. Effects of Agricultural Involvement on Food Consumption and Dietary Diversity

Explanatory Variables	1st Stage: Log share of on-farm income	2nd Stage: Log of Household Level Outcomes — Food and Calorie Consumption: Food Consumption Per capita	Consumption of Calories	Dietary Diversity: Food Consumption Score (FCS)	Simpson Index
	Effects of Share of On-Farm Income on Consumption and Dietary Diversity (2 Stage Least Square Estimates)				
Log share of on-farm income		0.293**	0.168**	0.102**	0.097*
Head Characteristics					
Sex of head (1=Male)	-0.113+	0.073**	0.030+	0.055**	0.062**
Age of head	-0.002	-0.002	0.003	-0.006**	-0.013**
Age of head squared	0.000	0.000	0.000	0.000**	0.000**
Highest education	-0.023**	0.013**	0.007**	0.008**	0.011**
Household Composition					
# of Kids 0-14 years	0.054**	-0.171**	-0.129**	-0.012**	-0.021**
# of Male adults 15-64 years	0.073*	-0.144**	-0.101**	-0.016*	-0.021*
# of Female adults 15-64 years	0.119**	-0.158**	-0.120**	-0.013+	-0.013
# of Individuals 65+ years	0.107	-0.214**	-0.156**	-0.046**	-0.080**
Agricultural Technology					
Used Seeds (D)	-0.145**	0.093**	0.050**	0.038**	0.043**
Use Inorganic Fertilizer (D)	1.086**	-0.221**	-0.098+	-0.077*	-0.017
Use Extension (D)	0.130*	0.032	0.056**	0.030**	0.021
Diversification and Credit Access					
Self-Employed (D)	-0.718**	0.326**	0.181**	0.134**	0.173**
Non-farm waged (D)	-0.910**	0.359**	0.207**	0.136**	0.178**
Farm waged (D)	-0.839**	0.168**	0.129**	0.032	0.035
Received credit (D)	-0.050	0.059*	0.013	0.032*	0.024
Region Fixed-Effects					
Rural Center (D)	-0.074	0.154**	0.021	0.002	-0.065**
Rural South (D)	0.006	-0.032	-0.022	-0.004	-0.009
Wealth and productivity factors					
Log Agro-Ecological potential	-0.016	0.005	-0.002	0.009*	0.009+
Household wealth index	-0.083**	0.133**	0.060**	0.065**	0.060**
Household agricultural asset index	0.101**	0.029**	0.023**	0.026**	0.051**
Share of sick adults chronically sick	0.021	0.048	0.047	-0.046*	0.046+
Mean Temperature - wettest quarter	-0.032**	0.008**	0.005**	0.003**	0.002
Household land holdings	-0.016	0.028**	0.016**	0.009**	0.004
Moderate nutrient avail. const. (D)	0.054	-0.044+	-0.016	-0.006	-0.034*
Severe nutrient avail. const. (D)	0.319**	0.000	-0.006	0.024	0.025
Instruments:					
Log # Dist. Ag. Officers/Household	0.235**				
Constant	7.154**	9.081**	7.050**	3.223**	-0.812**
Observations	8,872	8,872	8,872	8,872	8,872
R-Squared	0.173				
Endogeneity test (a)					
Wu-Hausman F (1;8844)		53.63	20.08	18.12	4.50
[p-value]		[0.000]	[0.000]	[0.000]	[0.034]
Weak-Identification test (b)					
Gragg-Donald Min. eigenvalue stat	31.05				
Crit. Val.:10% max IV size	16.38				

Note: (a) Ho: Share of on-farm income is exogenous; (b) Ho: Instruments are weak. Significance levels are: 1% (**), 5% (*), and 10% (+). Source: Malawi IHS3 Survey.

Table 6. Results of 3SLS System of Equations Estimation: Effects of Agricultural Involvement (Share of On-farm Income) on Share of Food Group Calorie Consumption

Explanatory Variables	3SLS System of Equations: Share of On-farm income and share of calories of food groups in total calories consumed (all shares ranging from 0 – 1)							
	Share of on-farm income	Share of Calories from Food Groups (shares across groups sum up to 1)						
		Grains	Roots	Pulses/ Nuts	Fruits/ Vegetables	Meat/Fish and Milk	Oils/Fats	Sugars
ENDOGENOUS:								
Share of on-farm income		0.670**	-0.258**	0.001	-0.065*	-0.187**	-0.234**	0.076
EXOGENOUS:								
Head Characteristics								
Sex of head (1=Male)	0.002	-0.009	-0.005+	0.001	-0.001	0.006*	0.007**	0.000
Age of head	0.002	0.003**	0.000	-0.001**	0.000	0.000	0.000	-0.001**
Age of head squared	0.000	0.000*	0.000	0.000**	0.000	0.000	0.000	0.000*
Highest education	-0.005**	0.001	-0.001*	0.000	0.000	-0.001**	0.000	0.001*
Household Composition								
# of Kids 0-14 years	0.004*	0.005**	0.000	-0.001**	0.000	-0.002**	0.002*	-0.003**
# of Male adults 15-64 years	0.006	-0.001	0.003	-0.001	-0.001	0.001	0.000	-0.001
# of Female adults 15-64 years	0.014**	-0.012*	0.005*	0.000	0.001	0.004*	0.003	0.000
# of Individuals 65+ years	0.006	0.009	-0.006	0.000	0.002	-0.005	0.002	-0.002
Agricultural Technology								
Used Seeds (D)	-0.012*	-0.009+	0.001	0.004**	0.001	-0.002*	0.002	0.002
Use Inorganic Fertilizer (D)	0.081**	-0.054**	0.015*	0.002	0.006*	0.011**	0.023**	-0.004
Use Extension (D)	0.027**	-0.037**	0.016**	0.002	0.004**	0.007**	0.009	-0.001
Diversification and Credit Access								
Self-Employed (D)	-0.373**	0.227**	-0.099**	0.003	-0.024*	-0.067**	-0.080**	0.037*
Non-farm waged (D)	-0.382**	0.242**	-0.105**	0.009	-0.023*	-0.072**	-0.088**	0.037*
Farm waged (D)	-0.247	0.191**	-0.066**	-0.003	-0.016*	-0.055**	-0.062**	0.010
Received credit (D)	-0.006	0.005	-0.002	0.000	0.000	-0.002	-0.003	0.001
Region Fixed-Effects								
Rural Center (D)	-0.003	0.168**	-0.077**	0.017**	0.004**	-0.117**	-0.005	0.004+
Rural South (D)	-0.047**	0.173**	-0.086**	0.020**	0.010**	-0.142**	-0.002	0.021**
Wealth and productivity factors								
Log Agro-Ecological potential	0.000	0.000	0.000*	0.000	0.000	0.000	0.000	0.000
Household wealth index	-0.024**	0.002	-0.005*	0.001	-0.002**	0.001	-0.003+	0.005**
Household agricultural asset index	0.018**	-0.020**	0.007**	0.001	0.001	0.004**	0.004**	0.002+
Share of sick adults chronically sick	-0.011	0.005	-0.002	0.005	0.005**	-0.005	-0.007*	-0.001
Mean Temperature - wettest quarter	-0.001**	0.000	0.000	0.000**	0.000**	0.001**	0.000	0.000
Household land holdings	0.015**	-0.004	-0.001	0.001	0.001	0.001	0.003*	-0.001
Moderate nutrient avail. const. (D)	-0.007	0.014*	-0.012**	0.001	-0.001	-0.004+	-0.003	0.006**
Severe nutrient avail. const. (D)	0.001	-0.030**	0.023**	-0.007**	0.000	0.007**	-0.007*	0.015**
Instrument:								
Log # Dist. Ag. Officers/Household	0.338**							
Constant	0.982**	-0.085	0.367**	0.097+	0.114**	0.169**	0.258**	0.041
Observations	8,684	8,684	8,684	8,684	8,684	8,684	8,684	8,684
Parameters	26	26	26	26	26	26	26	26
RMSE	0.234	0.248	0.116	0.068	0.039	0.080	0.091	0.066
R-Squared	0.500	-0.489	-0.275	0.023	-0.181	0.068	-0.488	-0.001
Chi-2	8668.550	917.370	643.540	204.790	190.240	3329.060	252.080	468.600
p-value	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000

Note: Significance levels are: 1% (**), 5% (*), and 10% (+).
Source: Malawi IHS3 Survey.

Annex Tables

Annex Table 1. Definition of Food Groups used in the Analysis

Food Groups	Food Items	Nutritional Attributes
Main Staples (Cereals and Tubers)	Cereals: Maize grain/flour; green maize; rice; finger millet; pearl millet; sorghum; wheat flour; bread; pasta; other Cereals. Roots/Tubers: Cassava tuber/flour; sweet potato; Irish potato; other tubers/plantain.	Energy dense, protein contents lower and poorer quality than legumes, micronutrients (bound by phytates).
Nuts and Pulses	Beans; pigeon pea; macadamia nut; groundnuts; ground beans; cow pea; other nut/pulse	Energy dense, high amounts of protein but of lower quality than meats, micronutrients (inhibited by phytates), low fat
Vegetables	Onion; Cabbage; Tanaposi; Nkhwani; Wild Green Leaves; Tomato; Cucumber; Other Vegetables/Leaves	Low energy, low protein, no fat, micronutrients
Fruits	Mango; Banana; Citrus; Pineapple; Papaya; Guava; Avocado; Apple; Other Fruits	Low energy, low protein, no fat, micronutrients
Meat, Fish, and animal products	Egg; Dried/Fresh/Smoked Fish (Excluding Fish Sauce/Powder); Beef; Goat Meat; Pork; Poultry; Other Meat	Highest quality protein, easily absorbable micronutrients (no phytates), energy dense, fat. Even in small quantities, improvements to the quality of diet are large.
Milk and Milk Products	Fresh/Powdered/Soured Milk; Yogurt; Cheese; Other Milk Product - Excluding Margarine/Butter or Small Amounts of Milk for Tea/Coffee	Highest quality protein, micronutrients, vitamin A, energy. However, milk could be consumed only in very small amounts and should then be treated as condiment. In such cases a reclassification is needed.
Sugar, sugar products and honey	Sugar; sugar cane; Honey; Jam; Jelly; Sweets/Candy/Chocolate; Other Sugar Products	Empty calories. Usually consumed in small quantities.
Oil and fats	Cooking Oil; Butter; Margarine; Other Fat/Oils	Energy dense but usually no other micronutrients. Usually consumed in small quantities.

Source: Adapted from WFP (2007), adjusted for Malawi IHS3.

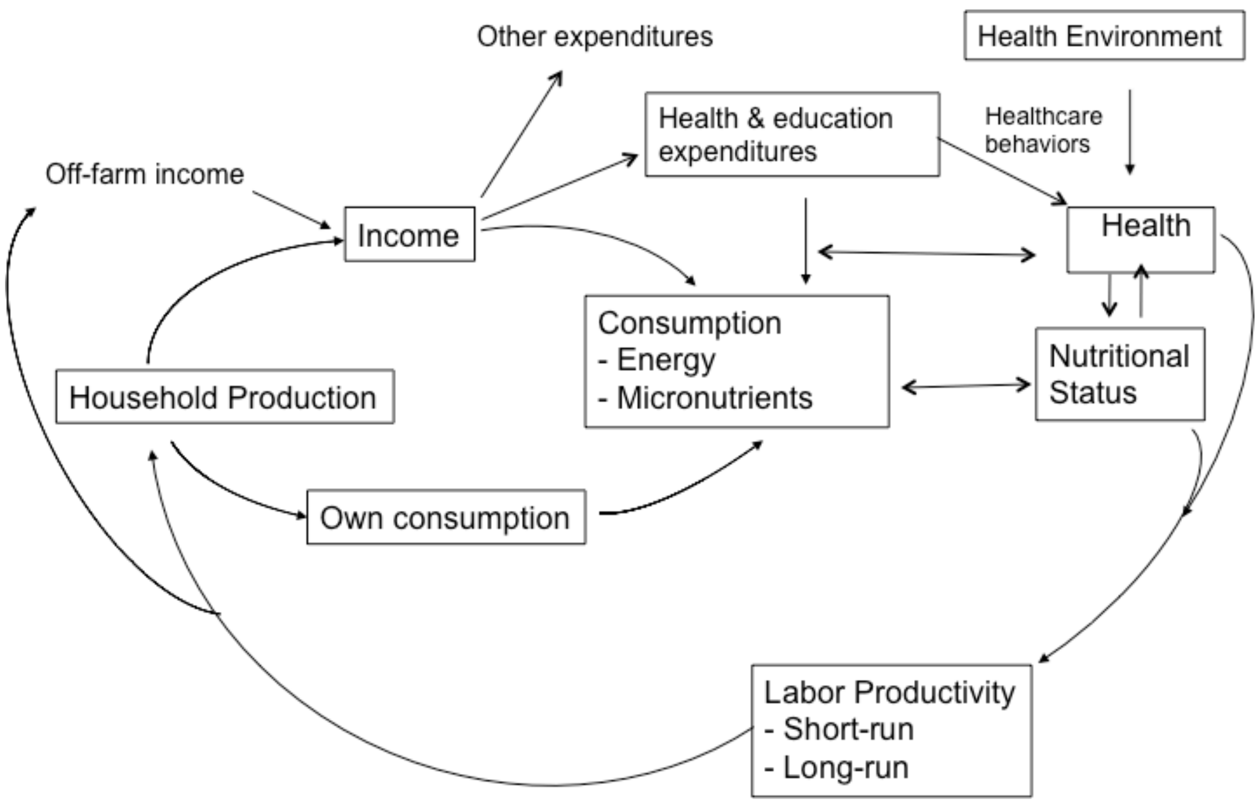
Annex Table 2. Food Consumption Inadequacy in Malawi

	Proportion of Households with Inadequate Food Consumption (%)[1]						
	All Households	Sex of the Head			Poverty Status		
		Male	Female	Difference	Non-Poor	Poor	Difference
Rural Malawi	22.7	20.3	29.9	9.6**	9.2	36.7	27.5**
Rural Region							
North	20.2	19.3	23.7	4.4	9.5	30.0	20.5**
Center	24.7	21.6	35.6	14.0**	10.1	43.6	33.5**
South	21.8	19.5	27.4	7.9**	8.1	33.7	25.6**

Notes: (1) Inadequate food consumption is defined as poor or borderline, i.e., a Food Consumption Score <35.
Significance level of the difference: 1% (**), 5% (*), and 10% (+).
Source: Malawi IHS3 Survey.

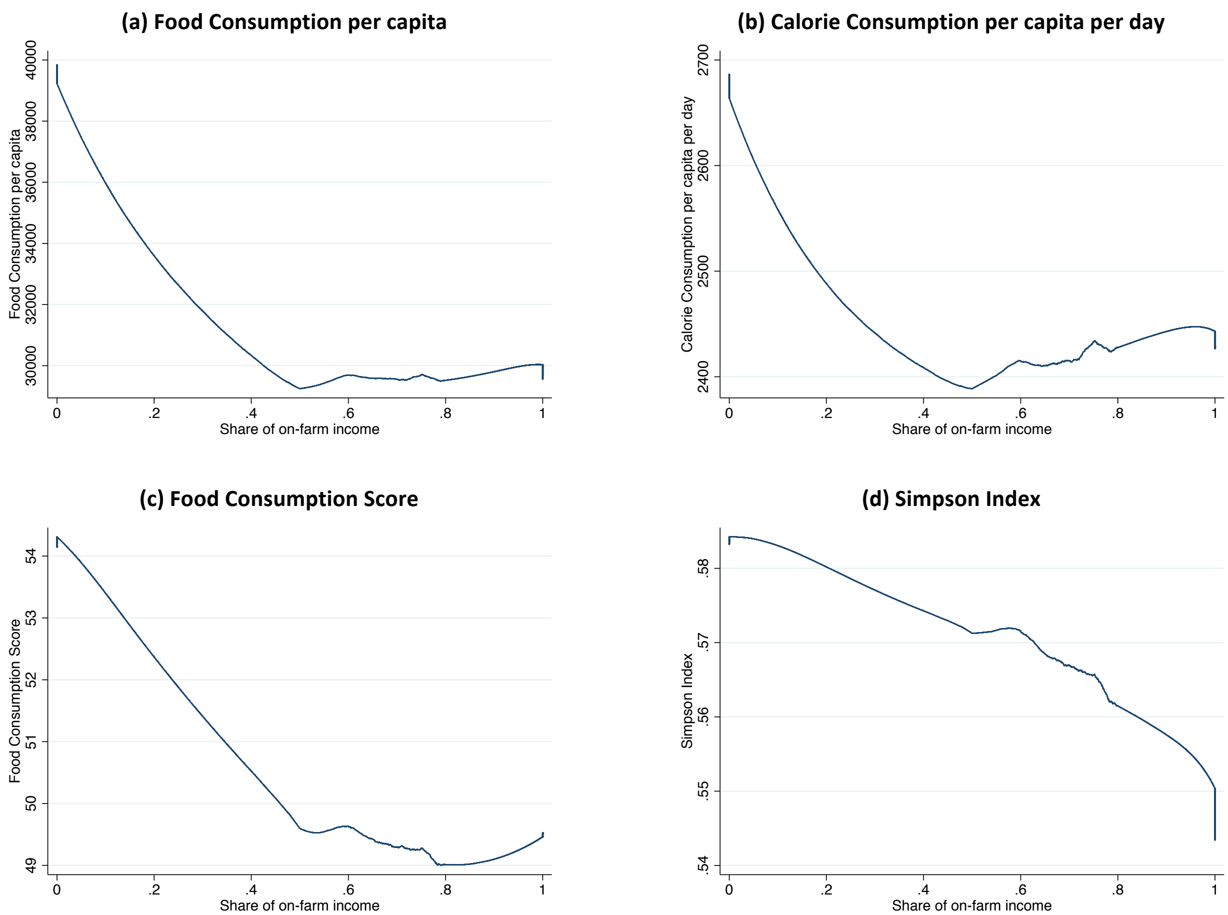
Figures

Figure 1. Agriculture-Nutrition Linkages Framework



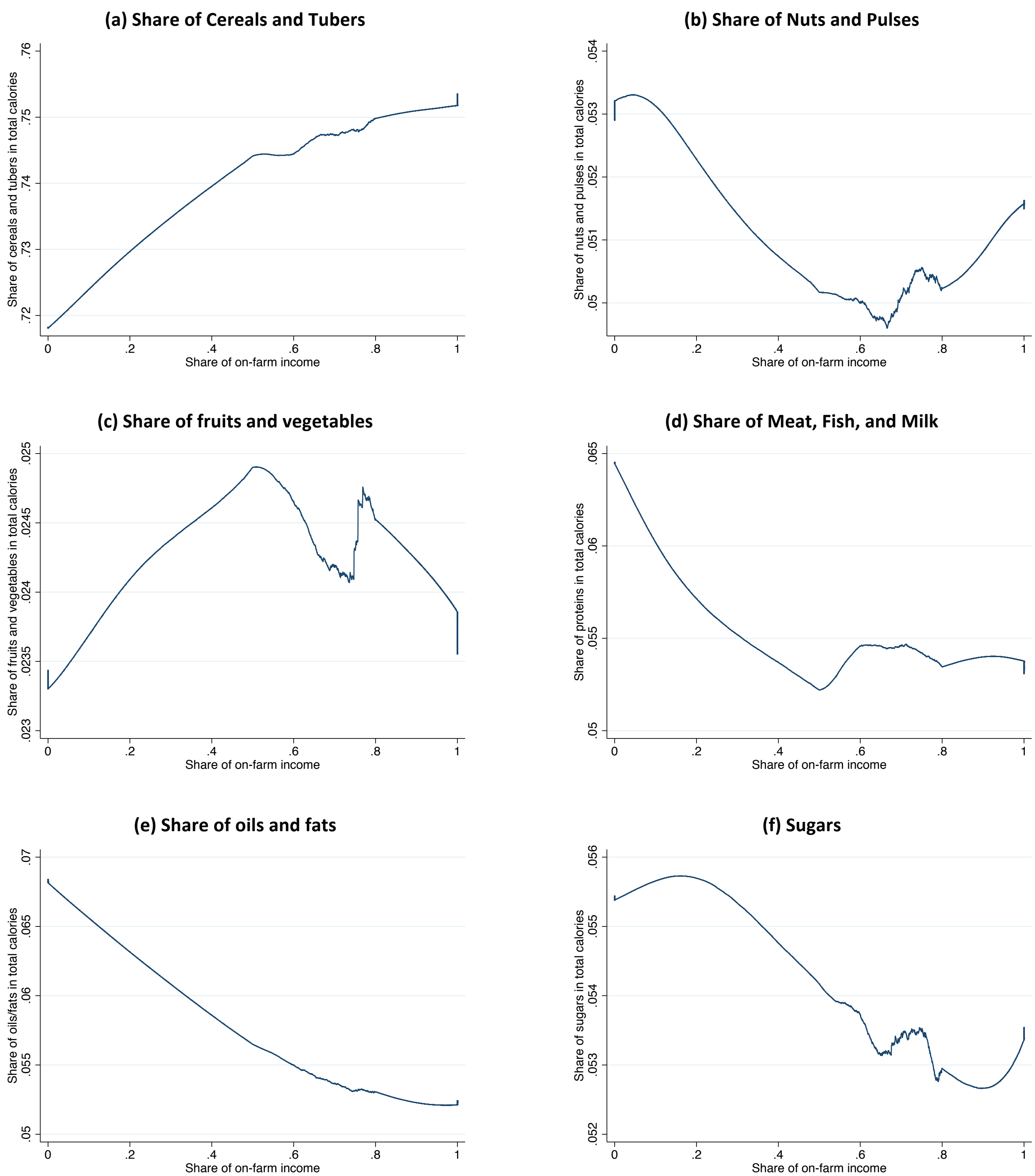
Source: Chung, 2012

Figure 2. Aggregate Food Consumption and Household Nutritional Outcomes by Share of On-farm income



Source: Author's computations with IHS3.

Figure 3. Share of Food Group Calorie Consumption, by Household Share of On-farm income



Source: Author's computations with IHS3.