Nutritional Effects of Agricultural Diversification and Commercialization in Children in Zambia

Selected Paper prepared for presentation at the Agricultural & Applied Economics Association's 2014 AAEA Annual Meeting, Minneapolis, MN, July 27-29, 2014.

Rhoda Mofya-Mukuka[1]

And

Christian H. Kulhgatz[2]

[1] Research Fellow, Indaba Agricultural Policy Research Institute, Lusaka, Zambia.

[2] Thünen Institute for Market Analysis, Braunschweig, Germany.

Abstract

Zambia and particularly the Eastern province have one of the highest rates of malnutrition in the world. The most vulnerable are the children from rural households which depend entirely on seasonal agricultural production and income, and survive on diets that are deficiency in proteins and other important nutrients. Agricultural diversification and commercialization provide alternative strategies for sustainable all-year-round household food and income availability. Applying Propensity Score Marching (PSM) and Generalized Propensity Score (GPS), this article evaluates the impact of agricultural diversification (in terms of calorie and protein production) and commercialization on reducing malnutrition in the Eastern province of Zambian. We use a uniquely rich dataset that comprises socioeconomic, agricultural and anthropometric data of 1120 children from five districts in the Eastern province. Results from PSM do not show significant impact of agricultural diversification and commercialization on reducing malnutrition while GPS results show that higher degrees of diversification reduce malnutrition. However, commercialization tends to have a negative effect particularly for short- and middle-term nutrition outcomes, where capital accumulation through higher purchasing power might have less impact. Policies need to consider the current diversification intensity of farmers and the different consequences on wasting and stunting when implementing diversification strategies. High levels of diversification could improve the wasting and underweight status of children by delivering a high amount of nutrients, but may come at the cost of reducing the efficiency of the farm and thus increasing the possibility of longer term stunting. Interventions focused on improving agricultural diversification and high degrees of commercialization may enhance adequate and diverse protein and calorie sources, while at the same time households will have excess produce for the market to meet their income demands.

Many people in Africa and particularly small children remain vulnerable to malnutrition and nutrient-related health problems. Studies indicate children that suffer from chronic malnutrition during the first two years of life to suffer from irreversible negative effects on brain and cognitive development (Unicef, 1990). This leads to reduced learning capacity in school and wage earning potential as adults.

Zambia has one of the highest rates of child malnutrition in the world. The most vulnerable are the rural households that highly depend on seasonal food production and survive on diets that are deficient in a variety of micronutrients. About 60.5% of the countries' population lives in the rural areas (CSO, 2010). According to the 2010 Zambia Living Conditions Monitoring Survey (LCMS, 2010), 46.7 % of the children in the country have stunted growth (z-score less than -2), 6% suffer from wasting and 13.3% are underweight. The malnutrition rates are even higher in the rural areas where 48.3% of the children are stunted, 14.6 % are underweight and 6.4% suffer from wasting. Although the prevalence of underweight children has declined from 25.1% in 1992 to 14.6% in 2010, it still remains a major concern as to whether Zambia will attain the Millennium Development Goals (MDG) target of 12.5 percent by 2015. Wasting cases are relatively moderate but worrisome as the rates have increased from 3.1% in 1996 to 6.4% 2010.

Agricultural diversification and commercialization provide alternative strategies for the rural households to improve diets (Hendrick and Msaki 2009; Khandker and Mahmud 2012); the former by yielding diverse food items for own consumption and the latter by increasing income and the household's ability to purchase a diverse range of food items. The growing of different groups of food crops contribute directly to a more diversified nutritional intake. At the same time, agricultural commercialization provide means of earning income that enable households to get goods and services (such as health) which become essential for sustaining their nutrition. There is overwhelming evidence in recent literature showing that increase in incomes during early childhood decreases stunting in the long-run (e.g Zere and McIntyre 2003; Monteiro et al. 2010; Alderman et al. 2006).

This article, evaluates agricultural diversification and commercialization as critical rural strategies for increasing access to nutritious foods in the eastern Province of Zambia. The article examines whether a diversified farm production significantly affects the nutritional status of children and also whether participation in agricultural markets improves their nutritional status.

The Eastern province is one of Zambia's high agricultural regions. It ranks third in terms of maize (the national staple food) and cattle production and first in the production of groundnuts, the main source of protein in the rural areas.

Table 1: Simpson Index of crop diversification per province.

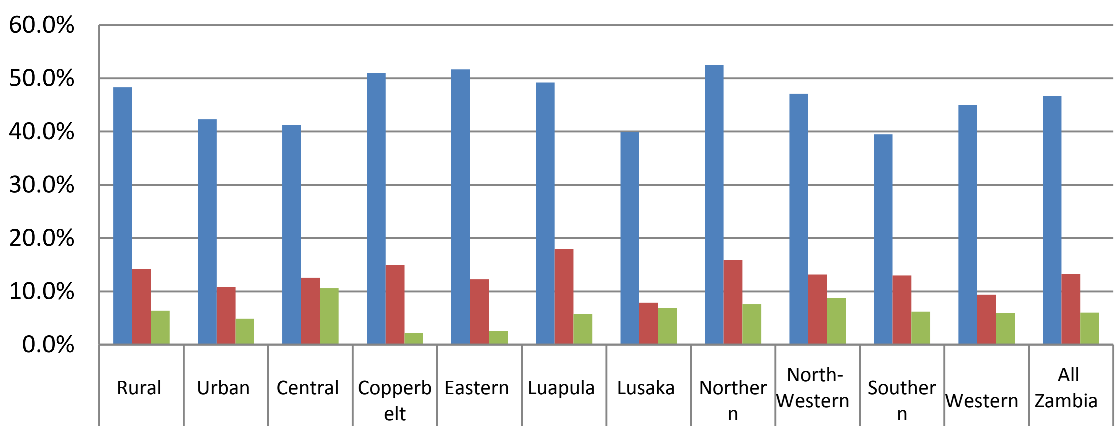
		Specialized	**Diversified**	
	Mean	Percentile 25	Median	Percentile 75
Central	0.41	0.2	0.48	0.61
Copperbelt	0.3	0	0.32	0.5
Eastern	0.47	0.38	0.5	0.63
Luapula	0.43	0.29	0.5	0.62
Lusaka	0.21	0	0.09	0.44
Muchinga	0.54	0.44	0.62	0.7
Northern	0.54	0.46	0.62	0.7
NorthWestern	0.4	0.23	0.46	0.58
Southern	0.31	0.09	0.33	0.5
Western	0.42	0.32	0.49	0.59
Zambia	**0.42**	**0.24**	**0.49**	**0.63**

In 2010/2011, the province produced 23% Maize, and 30% of groundnuts (IAPRI/CSO/MAL, 2012). As shown in table 1, Eastern province is also well known for high crop diversification.

The simpson index for crop diversification is 0.47, which is third highest out of ten provinces and above the national average of 0.42 (IAPRI/CSO, MAL, 2012).

Despite the high and diversified crop production, the protein and calories diversification is relatively low, less than 0.3 Simpson Index of diversification. This could explain the shocking high levels of child malnutrition recorded in the province. At 51.7%, the stunting rates are second highest in the country, higher than the national averages of 46.7%. Underweight rates stand at 12.3% while wasting rates are at 2.6% (Figure 1). The province also records high poverty rates of 80% which is the second highest in the country and remains above the national average of 75.5% (RALS, 2012). The high rate of nutrition amidst high and diversified agricultural production in the province is a paradox that requires evidence-based research drawing effective and sustainable solutions.

Figure 1: Incidence of stunting, underweight and wasting of children (3-59 months) by rural/urban and province, 2010, Zambia.



	Rural	Urban	Central	Copperbelt	Eastern	Luapula	Lusaka	Northern	North-Western	Southern	Western	All Zambia
Stunting	48.3%	42.3%	41.3%	51.0%	51.7%	49.2%	39.9%	52.5%	47.1%	39.5%	45.0%	46.7%
Underweight	14.2%	10.8%	12.6%	14.9%	12.3%	18.0%	7.9%	15.9%	13.2%	13.0%	9.4%	13.3%
Wasting	6.4%	4.9%	10.6%	2.2%	2.6%	5.8%	6.9%	7.6%	8.8%	6.2%	5.9%	6.0%

Source: Tembo and Sitko, 2013

Underlying factors of Stunting, Wasting and Underweight

Various factors have been associated with children nutritional status. Chiwele et. al (2010) argue that in Zambia malnutrition can occur even where supplies of food are sufficient, because of poor feeding practices, water and sanitation, HIV/AIDS and education. Household expenditure, location of the household (urban or rural) and level of education of the mother are other critical drivers of child nutritional status in Zambia (Masiye et. al, 2010). The level of education of the mother or any other care giver is another critical determinant of child malnutrition. In Ethiopia, women who received even a minimal education were found to be more aware of how to utilize available resources for the improvement of their own nutritional status and that of their families than those who had no education (Girma and Genebo, 2002).

The interactions between agricultural and health environments have implications on the utilization of food by the body (UNICEF, 1992). Without access to health services can lead to failure by the body to utilize the available food. At household level, the economic status of a household is an indicator of access to adequate food supplies, use of health services, availability of improved water sources, and sanitation facilities, are prime determinants of child and maternal nutritional status (UNICEF, 1992).

Evaluating the causal effect of diversification and commercialization

As indicated before, diversification as well as commercialization can potentially help improving the nutritional status of children. To quantify the effect of both measures, it is possible to employ the typical impact evaluation framework, in which diversification (commercialization) is seen as a treatment, and the nutritional status is the observed outcome. In the following section, we explain the econometric method by focusing on diversification as the treatment, but all the explanations also hold for commercialization.

In a first step, we use a simplified model in which treatment A is a binary variable, i.e. the farmer choses to diversify (A=1) or not (A=0). This is the conventional impact assessment scenario, and we will later on consider a more flexible approach. The expected treatment effect for the treated population is of primary significance. This effect is given as

$$\tau|_{A=1} = E(\tau \mid A=1) = E(O_1 \mid A=1) - E(O_0 \mid A=1) \tag{1}$$

where τ is the average treatment effect for the treated (ATT), A is a dummy for diversification decision, O_1 denotes the value of the outcome when the household diversified its production, and O_0 indicates the value of outcome in case the household did not diversify its production.

The measurement of the ATT is not trivial. The estimation problem arises due to the fact that it cannot be observed how a diversified household would have performed if it had not diversified its production, i.e. $E(O_0 \mid A=1)$ cannot be observed. Although the difference $[\tau^e = E(O_1 \mid A=1) - E(O_0 \mid A=0)]$ could be estimated, it would potentially be a biased estimator of the ATT, because the groups compared are likely to be different in their characteristics. This is because of self-selection of households, which is likely to occur when farm characteristics affect the utility that a farm derives from diversification or commercialization. To formalize the effect of farm characteristics on the treatment variable, we assume the following relationship between utility *U* and farm and household characteristics *Z* of farm household i:

$$U = \alpha' Z_i + \eta_i \tag{2}$$

where η_i indicates the residual. Given that the farmer maximizes utility by choosing whether to diversify or not to diversify, the probability of employing the diversification strategy is shown by the following equation:

$$\Pr(A_i = 1) = \Pr(U_{A,i} > U_{NA,i}) = \Pr(\eta_i > -\alpha' Z_i) = 1 - \Phi(-\alpha' Z_i) \quad (3)$$

Where $U_{A,i}$ is the maximum utility gained from choosing the treatment while $U_{NA,i}$ is the maximum utility derived from being in the control group. Φ indicates the distribution of the residual, which is logistic in the case of the Logit model we apply in our later analysis. Results of outcome comparisons between groups are biased even when farm characteristics are controlled for in simple regression analyses. To show this, consider a reduced-form relationship between the technology choice and the outcome variable such as

$$O_i = \alpha_0 + \alpha_1 A_i + \alpha_2 Z_i + \mu_i \quad (4)$$

Where O_i represents a vector of outcome variables for household i such as demand for inputs, A_i denotes a binary choice variable of diversification as defined above, Z_i represents farm level and household characteristics, and μ_i is an error term with $\mu_i \sim N(0, \sigma)$. The issue of selection bias arises if the error term of the technology choice η_i in equation (1) and the error term of the outcome specification μ_i in equation (2) are influenced by similar variables in Z_i. This results in a non-zero correlation between the two error terms, which would in turn lead to biased regression estimates when equation (4) is estimated with conventional OLS techniques. In particular, α_1 would not be a valid estimator of the ATT.

Several econometric techniques exist to re-establish a randomized setting in the case of self-selection. The difference-in-differences method is not applicable, as it requires panel data from several time periods, which is not provided by RALS data. The instrumental variables technique relies on parametric assumptions regarding the functional form of the relationship between the outcomes and predictors of the outcome, as well as on the exogeneity of the instruments used. Since this approach is quite sensitive to violations of these strict assumptions, we follow the matching approach, in which households of the group of diversified farmers are matched to households in the control group that are similar in their observable characteristics.

Propensity Score Matching Approach

Given the multitude of factors potentially influencing the adoption decision, it is hardly possible to match each household of the group of adopters with an adequately similar household in the group of non-adopters. As a solution to this problem, Rosenbaum and Rubin (1983) have shown that it is possible to use the propensity of adoption as a single indicator for similarity, making it a balancing score in the matching process. The propensity score is defined as the conditional probability that a farmer is diversified , given pre-adoption characteristics (Rosenbaum and Rubin, 1983). To create the condition of a randomized experiment, the PSM employs the unconfoundedness assumption also known as conditional independence assumption (CIA), which implies that once X is controlled for, participation is random and uncorrelated with the outcome variables. The PSM can be expressed as,

$$p(X) = \Pr\{P = 1 | X\} = E\{P | X\} \qquad (5)$$

where P = {0,1} is the indicator for being diversified and X is the vector of household and farm characteristics. Given CIA, the conditional distribution of X, given p(X) is similar in both groups of participation and non participation, so the effect measured after balancing with the propensity-score is like in a randomized experiment.

The CIA is a strong assumption. In case selection into treatment is based on unmeasured characteristics, there may still be systematic differences between outcomes of diversified and non-diversified households even after conditioning on the propensity score (Smith and Todd, 2005). However, Jalan and Ravallion (2003) pointed out that the CIA assumption is no more restrictive that those of the IV approach employed in cross-sectional data analysis.

In our study, we match on the odds ratio, since Leuven and Sianesi (2003) indicated it to be the general suggestion for household survey data. These odds ratio is calculated with a Logit model of equation (3). The empirical analysis is then carried out by employing the approach suggested by Rosenbaum and Rubin (1983).

After having computed the balancing score for each household, the average treatment effect for the treated (ATT) is estimated by the average differences of matched pairs with similar score values. This can be stated as

$$\begin{aligned}\tau &= E\{O_1 - O_0 \mid A=1, p(X)\} \\ &= E\{E\{O_1 \mid A=1, p(X)\} - E\{O_0 \mid A=0, p(X)\} \mid A=1\}\end{aligned} \tag{6}$$

Several techniques have been developed to match adopters with non-adopters. In the current paper the Nearest Neighbour Matching (NNM) method is employed.

Treating Diversification and Commercialization as continuous variables

The previous methods treated diversification and commercialization as a binary decision variable. This is probably be an oversimplification, since households produce at different intensity levels of diversification and commercialization. These various levels may have different effects on the nutritional status. In a final step, we change this econometric setup, and measure the impact of different levels of diversification and commercialization. For this, we use the method proposed by Hirano and Imbens (2004) and employ the generalized propensity score to balance the differences among farms of different intensity levels. The unbiased heterogeneous impact of different intensities of diversification and commercialization on health outcomes can then be illustrated with dose response functions.

For each household i, we observe the vector of pre-treatment variables X_i, the actual level of treatment received, T_i, and the outcome variable associated with this treatment level $O_i = O_i(T_i)$. Of interest is the dose response function (DRF), which relates to each possible production intensity level t_i, the potential welfare outcome $O(t)$ of farm household i:

$$\theta(t) = \mathrm{E}[O_i(t)] \quad \forall \;\; t \in \mathcal{T} \;\; \text{where} \;\; \mathcal{T} = (0, \dots, 1] \tag{7}$$

where θ represents the DRF, and t is the treatment level, which is measured as a diversification index (the Simpson index) or as the share of crops sold in total crop revenues (commercialization index). Similar to the CIA assumption in the PSM setting for dichotomous treatment variables, we presume weak unconfoundedness.[1] In order to adjust for a large number of observable characteristics, Hirano and Imbens (2004) suggest estimating the generalized propensity score

(GPS), which is defined as the conditional density of the actual treatment given the observed covariates. Formally, let $r(t,x) = f_{T|X}(t|x)$ be the conditional density of potential treatment levels given specific covariates. Then the GPS of a household i is given as$R_i = r(T_i, X)$. The GPS is a balancing score, i.e. within strata with the same value of $r(t,X)$, the probability that $T = t$ does not depend on the covariates X_i. Due to its balancing property, the GPS can be used to derive unbiased estimates of the DRF (Hirano and Imbens, 2004). For this, the conditional expectation of the outcome first needs to be calculated as $\gamma(t,r) = E[O_i|T_i = t, R_i = r]$. The average DRF of equation (5) can then be estimated at particular levels of treatment as follows:

$$\theta(t) = \mathrm{E}\left[\gamma\left(t, r(t, X_i)\right)\right] \tag{8}$$

The GPS is estimated with a generalized linear model (GLM) with covariates X_i and a fractional logit (Flogit) specification, which takes into account that both of the analyzed treatment variables (diversification and commercialization) range between 0 and 1.[2]

The common support condition is imposed by applying the method suggested by Flores et al. (2009).[3] We test the balancing property of the estimated GPS by employing the method proposed by Kluve et al. (2012). The conditional expectation of the outcome for each farm is estimated using a flexible polynomial function, with quadratic approximations of the treatment variable and the GPS, and interaction terms (Hirano and Imbens, 2004). The specification is estimated using OLS regression for continuous welfare outcomes. Then the DRF of equation (6) is evaluated at 10 evenly distributed levels of agricultural diversification or commercialization. Confidence bounds at 95% level are estimated using the bootstrapping procedure with 1000 replications.

Data

We use a uniquely rich dataset that comprises socioeconomic, agricultural and anthropometric data. The study covers 1120 children from the Eastern province of Zambia with data collected in two rounds. The first dataset is from the 2012 Rural Agricultural Livelihood Survey (RALS), a nationally representative dataset covering 8839 households. The RALS, which was conducted by the Indaba Agricultural Policy Research Institute (IAPRI) in partnership with the Central Statistical Office (CSO) and the Ministry of Agriculture and Livestock, provides information for calculating crop diversification and agricultural commercialization.

The second dataset is Anthropometric data collected from the same households and is used to calculate stunting, wasting and underweight in children. This dataset also provides variables related to the health environment. The data was collected in December 2012 which gives almost two years from January 2011 when the household begin to consume the produce from the 2010/11 farming season, to the time of collection of Anthropometric data. This period was very important to examine height-for-age cumulative effects of past nutrition deprivations. The Anthropometric data included only children (0 – 59 months) from the 1120 households in five districts in Eastern province.

We calculate diversification using the Simpson Index over production of major food groups; starchy foods, legumes-nuts-seeds, starchy vegetables, non-starchy vegetables, starchy fruits, non-starchy fruits, dairy, and eggs. Meat and meat products could not be added to the list because these were consumed very rarely. We measure production in two ways; firstly in terms of calorie production (CDIV), and secondly in terms of protein production (PDIV).

$$PDIV = 1 - \sum_{i=1}^{S} p_i^2$$

$$CDIV = 1 - \sum_{i=1}^{S} c_i^2$$

where S is the number of food groups. Commercialization was measured as an index derived from the share of agricultural sales in household's total value of agricultural production.

For propensity score matching, we use the median of CDIV, PDIV and COM to distinguish between treated and untreated farm households. Farm characteristics that we control for with the propensity-score include gender, education, household composition, wealth, land tenure, remittances, maize subsidy receipt, infrastructure and location. Variables measuring these characteristics are included in the Logit model for the conventional propensity-score matching approach, and the GLM model for the generalized propensity-score approach. Descriptive statistics for treatment, outcome, and balancing variables are found in Table 2.

Table 2. Descriptive statistics of the variables.

Variable	Description	Mean	Std. Dev.
Nutritional outcome variables			
haz	Length/height-for-age z-score	-1.86	1.69
waz	Weight-for-age z-score	-0.86	1.18
whz	Weight-for-length/height z-score	0.26	1.51
Treatment variables			
Calorie Simpson Index	index "=1-sum of squared calorie shares of the produce.	0.26	0.19
Protein Simpson Index	index "=1-sum of squared crop protein shares of the produce.	0.28	0.18
Commercialization	household commercialization index	0.50	0.27
Farm characteristics			
FHHdefacto	=1 if de facto female-headed HH	0.12	0.33
noformaled	=1 if HH head has no formal education	0.18	0.39
grade1_4	=1 if HH head completed lower primary (grades 1 to 4)	0.18	0.39
grade5_7	=1 if HH head completed upper primary (grades 5 to 7)	0.34	0.47
agehead	Age of the HH head	40.48	12.51
ftesum	Full-time equivalent HH members	6.19	2.57
shareAgeun~5	Share of household members aged below 5	0.20	0.14
shareAge5_14	Share of household members aged 5 to 14	0.30	0.19
shareAbove60	Share of household members aged 60	0.04	0.12
deathinfam~y	=1 if the household experienced death of a member within the reference perion	0.05	0.23
landholdsz12	Total land holding size less rented in and borrowed in	3.58	3.09
landother	sum of land borrowed in and rented in	0.16	0.81
Landtitled	land with title deeds	0.28	1.56
deflstock	Value of livestock (real ZMK, 2007/08=100) [3]	2781176.00	4534321.00
defvalequip	Value of farm equipment (ZMK/10,000; 2007/08=100)	43.07	88.94
fisphh	=1 if HH acquired FISP fertilizer	0.47	0.50
remit_c	Cash remittances received	139725.90	808848.70
remit_m	Value of maize received	7527.23	32657.21
remit_o	Value of other commodities received	15975.00	110869.80
bomai	Km from the homestead to the nearest boma	31.20	20.74
feedroadi	Km from the homestead to the nearest feeder road	1.81	5.07
agrodealeri	Km from the homestead to the nearest agro-dealer	24.99	20.84
clinic_max	distance to the nearest clinic	6.49	5.97
district2	dist==Katete	0.22	0.42
district3	dist==Lundazi	0.25	0.43
district4	dist==Nyimba	0.10	0.30
district5	dist==Petauke	0.19	0.39

Notes: We included interaction terms between boma distance and each district dummy as well as quadratic terms for age of household head and land holding size. Reference for district dummy variables is the Chipata district and for the education dummy variables education grades above 7.

[3] At the time of the RALS, the Kwacha-dollar rate was $1 = ZMK5012.

Results

a. Propensity Score Matching

We used the variables of Table 2 to calculate the odd ratios with a logit model. For the sake of brevity, we do not show the results of the propensity-score calculation, since this model is not for interpretational purposes but just for deriving a sample of matched households that are well balanced in their characteristics.[4] The estimation results of the PSM method are shown in Tables 3, 4 and 5.

Table 4 shows the result from treatment with calorie diversification index. These results indicate that, in terms of stunting, underweight and wasting levels, there is no significant difference between children from households which were above or equal to 0.23 calorie production diversification (Simpson index) and those from less than 0.23 Simpson Index. These results are consistent with the general summary of the meta-analysis of Masset et al. (2012).

Table 4: ATT with calorie diversification index as treatment variable

Variable	Sample	Treated	Controls	Difference	S.E.	T-stat
waz	Unmatched	-.8609319	-.862720721	.001788821	.070863297	0.03
	ATT	-.8609319	-.913154122	.052222222	.12039295	0.43
haz	Unmatched	-1.84740143	-1.81327928	-.034122154	.097209559	-0.35
	ATT	-1.84740143	-1.93446237	.087060932	.155240961	0.56
whz	Unmatched	.292903226	.217063063	.075840163	.090368959	0.84
	ATT	.292903226	.225716846	.06718638	.144038334	0.47

Note: S.E. does not take into account that the propensity score is estimated.

Table 5 shows results from treatment with protein production diversification Simpson index. Unlike the treatment with calorie production diversification, the treatment with protein diversification has significant impact on reducing stunting (by 0.30 HAZ), but not on reducing wasting and underweight. The impact is positive in that children from more protein production diversified households (0.26 and above Simpson Index), have a higher HAZ than those from less diversified households. However, the difference of 0.30 is too small if children are severely stunted. As mentioned above, PSM does not account for different intensity levels of protein production diversification, such that child nutritional status is affected differently at different intensities.

Table 5: ATT with protein diversification index as treatment variable

Variable	Sample	Treated	Controls	Difference	S.E.	T-stat
waz	Unmatched	-.812468694	-.911624549	.099155855	.070801306	1.40
	ATT	-.810752688	-.877741935	.066989247	.10514265	0.64
haz	Unmatched	-1.79872987	-1.86232852	.063598645	.09719685	0.65
	ATT	-1.79587814	-2.10082437	.304946237	.139056961	2.19
whz	Unmatched	.323255814	.186299639	.136956175	.090304752	1.52
	ATT	.323405018	.420125448	-.09672043	.132186981	-0.73

Table 6 shows the results when treatment is commercialization index. The results show no significant impact on all the three malnutrition measures.

Table 6: ATT with commercialization index as treatment variable

Variable	Sample	Treated	Controls	Difference	S.E.	T-stat
waz	Unmatched	-.85099278	-.87255814	.02156536	.070860821	0.30
	ATT	-.861797753	-.937827715	.076029963	.126946814	0.60
haz	Unmatched	-1.76144404	-1.89871199	.137267942	.097128309	1.41
	ATT	-1.75303371	-1.93011236	.177078652	.18571376	0.95
whz	Unmatched	.200794224	.308890877	-.108096653	.09033999	-1.20
	ATT	.178651685	.298370787	-.119719101	.183225724	-0.65

Note: S.E. does not take into account that the propensity score is estimated.

The insignificance of all these measures could however indicate that the application of the PSM method is not appropriate, as it requires the somewhat arbitrary categorization in diversified and non-diversified farmers and commercialized and non-commercialized farmers, respectively. Households however produce at different intensity levels of diversification and commercialization, which could have different effects on child nutrition. Since PSM cannot capture such heterogeneous effects of different intensity levels, we employ the GPS approach in the following section.

b. Generalized Propensity-Score

In the following section, results of the Generalized Propensity Score approach are presented. As with the PSM approach, we do not report the GLM estimation results due to limitation in space. The balancing property is tested by first regressing the treatment variable on each of the variables included in the GLM, and then replicating these regressions but with the inclusion of the estimated GPS as an additional explanatory variable. If the GPS successfully balances for the different farm characteristics, the inclusion of it should render any significant effect between the treatment variable and the farm characteristics variable insignificant.

For the calorie index, 6 variables are significant at the 1% level before the GPS is included. After the GPS was introduced into all regressions, there is no variable with significant effect on the treatment intensity anymore. In case of the protein index and before the incorporation of the GPS into the regression, 7 variables were significant at 1% level, 2 were significant at 5% level and 1 was significant at 1% level. After the inclusion of the GPS in the PDIV equations, one variable remains significant, however at a low 10% significance level. For commercialization, the test shows that before the inclusion of GPS, 6 variables are significant at 1% level and 4 are significant at 10% level, while none is significant when the GPS is included. We therefore conclude that the variables used for balancing fairly well balance the differences in farm characteristics and go on with the analysis of the treatment effect.

i. Treatment with calorie production diversification (Simpson Index)

Figure 1 depicts the effect of different intensities of calorie diversification on the nutritional status of children. In each of the three diagrams, the x-axis indicates the intensity of calorie diversification measured in terms of the Simpson index (CDIV), and the y-axis measures the expected effect on a) HAZ, b) WAZ and c) WHZ at the given level of diversification. Once the continuous nature of diversification is taken into account, trends can be seen how calorie diversification affects the nutritional status of children.

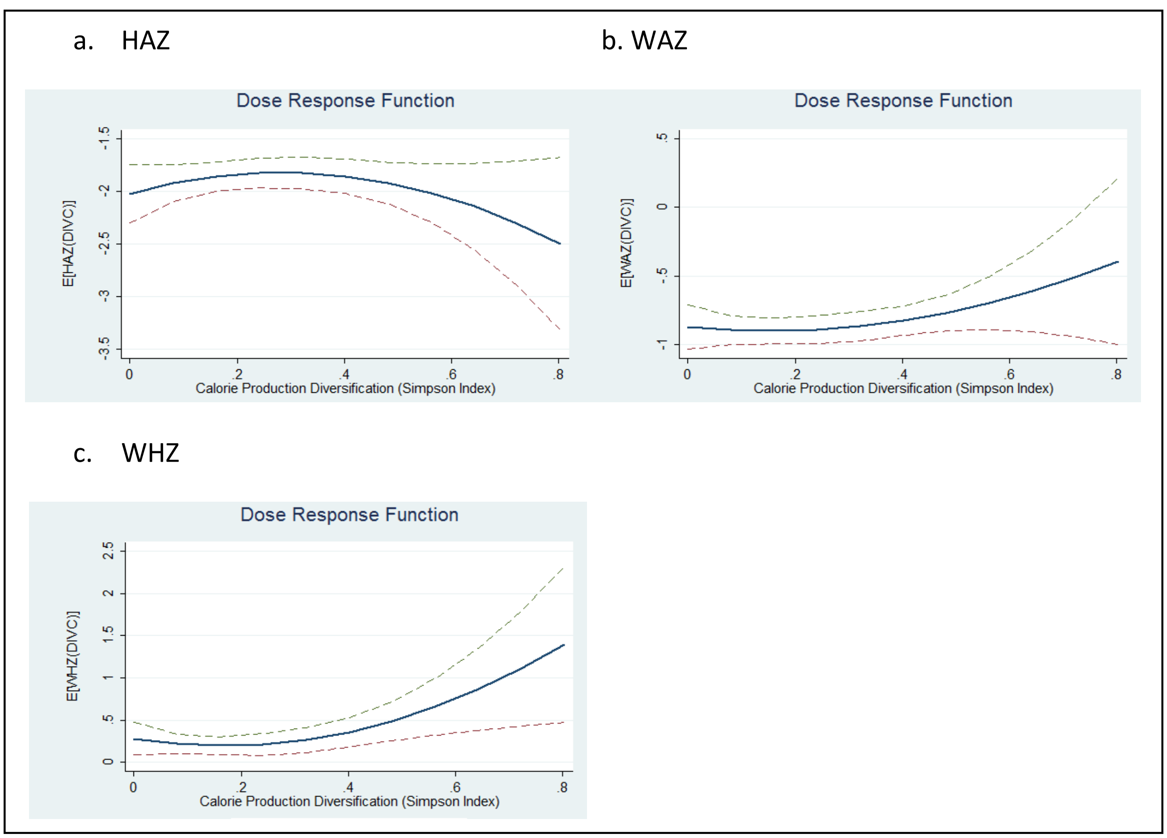
The HAZ indicator shows that the long term nutritional effect of calorie diversification tends to be positive at low diversification levels (i.e. high specialization), however at a relatively marginal rate. The DRF has a maximum at roughly 0.3, and becomes negative at high levels of diversification. An explanation for this non-linear relationship might be that on the one hand, specialization in very few crops results in a permanently less diverse diet with quickly arising long-term consequences for nutritional status of the child. On the other hand, extremely high diversification levels could reduce food security of children due to a less efficient production structure that delivers fewer amounts of nutrients than less diversified farms could produce.

Given that the median of calorie diversification is at 0.23, it could be concluded that the calorie production do not need any further diversification or specialization.

The DRFs for the effect of calorie diversification on WAZ and WHZ are similar, but show a very different shape than the HAZ function. Both graphs show a positive relationship between calorie diversification and the children's nutritional status. High levels of diversifications may prevent households from short term shock situations due to their stable provision of diverse set of nutrients that are correlated with calories from different agricultural products.

The spread in the confidence intervals at high levels above 0.5 is a typical pattern in all of our study's graphs on diversification. This indicates that there are few farms that produce at these high levels of diversification, so although the average effect has a clear trend, statistical predictions become shakier.

Figure 1: Treatment with Calorie Production Diversification (Simpson Index)

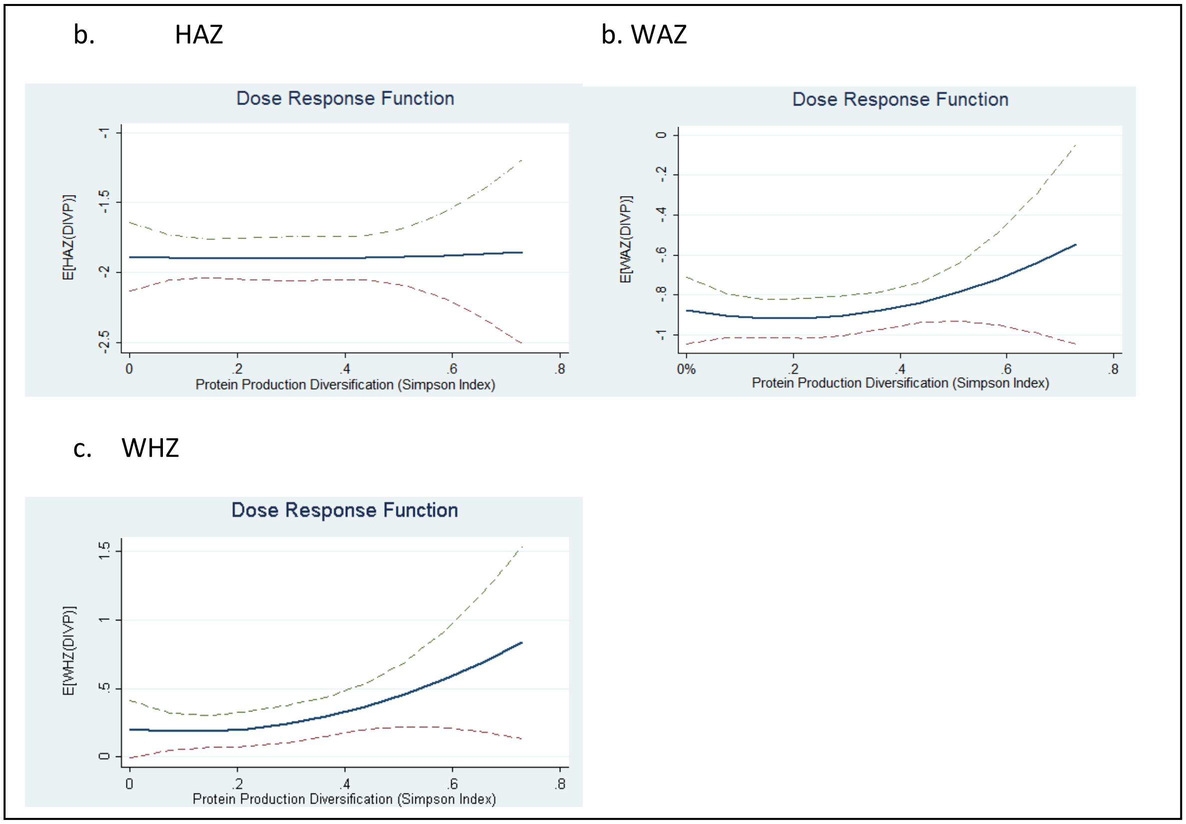


Note: the straight line is the dose response function and dashed lines indicate the 95% confidence interval.

ii) Treatment with protein production diversification (Simpson Index)

Figure 2 presents the heterogeneous effect of protein diversification on the nutritional outcomes. The effects are very similar to the calorie diversification, with one major difference. The HAZ dose response function remains flat over the whole range of treatment levels, therefore indicating that for stunting levels there are in fact no significant effects to expect from a diversification in protein sources. This is not surprising, given that the data used for calculating the treatment variable did not provide enough timeframe to establish impact on long-term nutritional status. However, the protein effect on WAZ and WHZ are clearly positive and significant at quite high levels of diversification.

Figure 2: Treatment with Protein Production Diversification (Simpson Index)

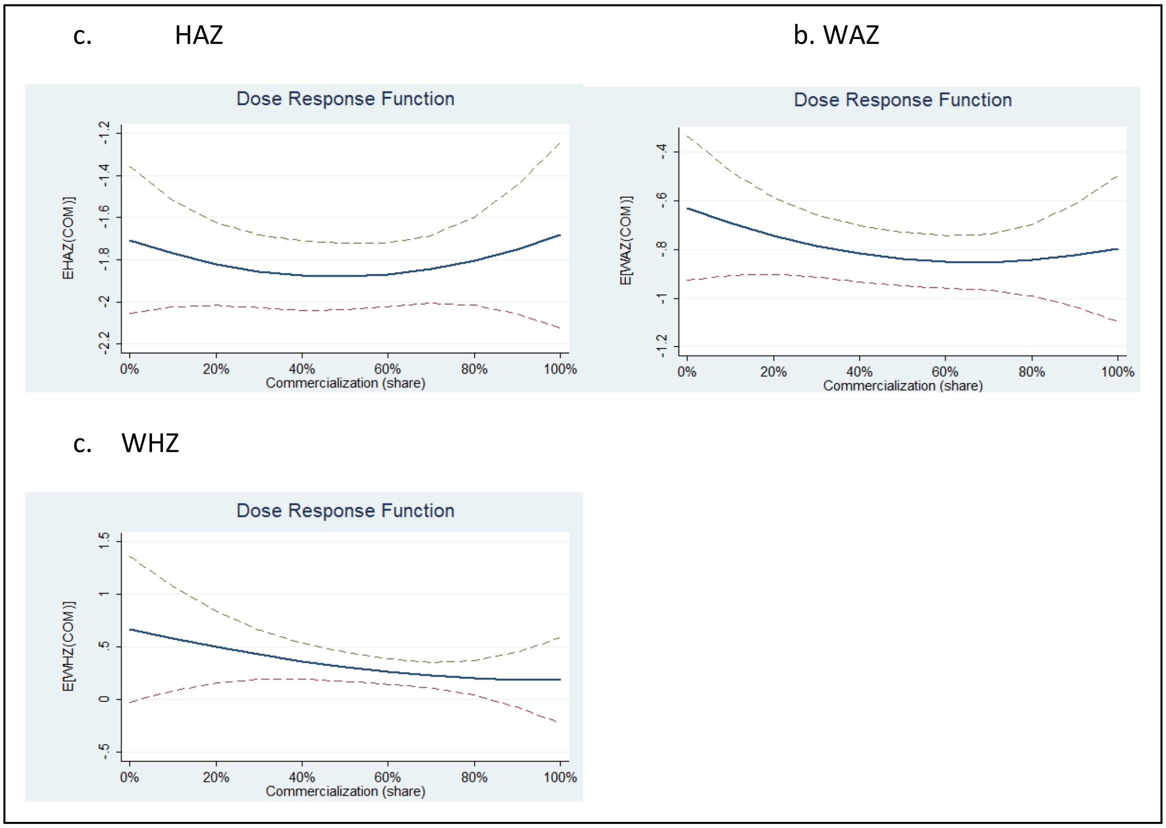


Note: the straight line is the dose response function and dashed lines indicate the 95% confidence interval.

iii. Treatment with Agricultural Commercialization Index

Figure 2 presents the effect of commercialization on the nutritional outcomes. All indicators show a similar pattern. Commercialization seems to have a negative slope for WAZ and WHZ, and also for most intensity levels of HAZ. However, at higher intensities of commercialization, commercialization seems to become more beneficial for the nutritional long-term status, but it only reaches similar levels as those hoseholds with no commercialization at all. There might be two strategies to tackle the large problem of stunting in Zambia, either specializing in cash crops, or specializing in a subsistence farm, which maybe has other income sources than agriculture.

Figure 3: Treatment with Agricultural Commercialization Index



Note: the straight line is the dose response function and dashed lines indicate the 95% confidence interval.

Conclusion and Policy Recommendations

Agricultural diversification and commercialization remain critical for improving the nutrition status of children. However, there are important aspects of improving nutritional status of children with the two agricultural strategies that need to be taken into account. First, the above results have shown that intensity of treatment at household level matters in the nutrition status of the children. Very high levels of diversification can improve nutritional status while smaller levels do not have significant impacts. Second, it is important to strategies agricultural production diversification according to the food groups because different food groups have varying impact on different forms of malnutrition. The impact of protein production diversification is positive and significant at high levels of diversification for short and medium term malnutrition effects. However, the impact on long term malnutrition is not significant even with increasing intensity of diversification. On the other hand, the impact of calorie diversification is non-linear, an indication that specialization in very few crops results in a permanently less diverse diet with quickly arising long-term consequences for nutritional status of the child. This is consistence with food production and consumption patterns in rural Zambia which is mainly based on calorie consumption. These results explain why stunting is high despite a diversified calorie production.

Third, commercialization has a significant but negative effect on improving both the long and the short-term malnutrition status in Children. Given the high commercialization index of 0.5, the results imply that most households sell most of their agricultural produce, regardless of the quantities produced, leaving very little for home consumption. It can further be concluded that the revenue realized from the sales, is not spent on purchase of nutritious food.

Comparing the results from PSM and GPS, it is evident that GPS provides more consistent results and explains the impact at different intensities. The results, in part, explain why malnutrition levels have remained very high in the province despite the fact that the agricultural production high. Policies need to consider the current diversification intensity of households and the different consequences on wasting and stunting when implementing diversification strategies. High levels of diversification could improve the wasting and underweight status of children by delivering a high amount of nutrients, but may come at the cost of reducing the production efficiency of the households and thus increasing the possibility of longer term stunting. Interventions focused on improving agricultural diversification and high degrees of commercialization may enhance adequate and diverse protein and calorie sources, while at the same time households will have excess produce for the market to meet their income demands.

References

Augustol C. M et al. 2010. Narrowing socioeconomic inequality in child stunting: the Brazilian experience, 1974-2007. *Bulletin of the World Health Organization*, *88*(4): 305-311.

Alderman, H., H. Hoogeveen and M. Rossi .2006. Reducing child malnutrition in Tanzania: Combined effects of income growth and program interventions. *Economics & HumanBiology* 4(1): 1–23.

CSO, 2010. 2010 Census of Population National Analytical Report. Central Statistical Office 2012. Lusaka.

Flores, Carlos A. and Alfonso Flores-Lagunes. 2009. Identification and Estimation of Causal Mechanisms and Net Effects of a Treatment under Unconfoundedness. University of Miami, mimeo.

Hendrick, S, L and Msaki, M.M. 2009. The impact of smallholder commercialization of organic crops on food consumption patterns, dietary diversity and consumption elasticities. Agrekon 48: 2. Accessed October 2013. Available at http://ageconsearch.umn.edu/bitstream/53383/2/5.%20Hendriks%20&%20Msaki.pdf.

Hirano, K., and G. W. Imbens. 2004. The propensity score with continuous treatments. In *Applied Bayesian Modeling and Causal Inference from Incomplete-Data Perspectives*, ed. A. Gelman and X.-L. Meng, 73–84. West Sussex, England: Wiley InterScience.

IAPRI/CSO/MAL. 2012. 2012 Rural Agricultural Livelihoods Survey Data (RALS). IAPRI. Lusaka.

Khandker, S.R and W. Mahmud (2012). *Seasonal Hunger and Public Policies*: Book. World Bank. Washington DC, USA.

Kluve, J., H. Schneider, A. Uhlendorff and Z. Zhao. 2012. Evaluating continuous training programs using the Generalized Propensity Score. Journal of the Royal Statistical Society, Series A 175, 587–617 (2012).

Leuven, E. and B. Sianesi .2003. "PSMATCH2: Stata module to perform full Mahalanobis and propensity score matching, common support graphing, and covariate imbalance testing". http://ideas.repec.org/c/boc/bocode/s432001.html. This version 4.0.8.

Masset, E., L. Haddad, A. Conelius and Jairo Isaza-Castro. 2012. Effectiveness of Agricultural Interventions that aim to Improve Nutritional Status of Children: Systematic Review. *BMJ 2012; 344 doi: http://dx.doi.org/10.1136/bmj.d8222*

Rosenbaum, P. R. and D.B Rubin. 1983. The Central Role of the Propensity Score in Observational Studies for Causal Effects. *Biometrika* 70 (1): 41–55. doi:10.1093/biomet/70.1.41

UNICEF. 1992. Strategy for Improved Nutrition of Children and Women in Developing Countries. A UNICEF Policy Review. UNICEF, New York.

Zere, E. and D. McIntyre. 2003. Inequities in under-five child malnutrition in South Africa. *International Journal for Equity and Health* 2:7.

---

[1] This assumption essentially postulates that, once all observable characteristics are controlled for, there is no systematic selection into specific levels of diversification / commercialization intensity left that is based on unobservable characteristics (Flores et al., 2009).

[2] The fractional logit model is implemented as a GLM with Bernoulli distribution and a logit link-function.

[3] We thank Helmut Fryges and Joachim Wagner for granting us access to a modified Stata program that allows the imposition of common support..

[4] The results can be obtained from the authors by request.