Djokoto *et al. Agric & Food Secur (2017) 6:6*
DOI 10.1186/s40066-016-0082-4

Agriculture & Food Security

RESEARCH

Open Access

Vegetable diversification in cocoa-based farming systems Ghana

Justice G. Djokoto[1*], Victor Afari-Sefa[2] and Albert Addo-Quaye[3]

Abstract

Background: As part of dynamic livelihood coping strategies, some farmers in Ghana's cocoa belt have diversified away from traditional cocoa production to other high-value crops including vegetables, to the extent of diversifying within vegetables. This study assessed the extent of diversification of vegetables among farmers in Ghana's cocoa belt and determined the factors that explain the variability in the diversification indices. A small-sample-size formula (http://www.surveysystem.com/sscalc.htm) that was based on an estimated population of the sample was used to arrive at 621 farmer respondents from the Ashanti and Western Regions of Ghana. A combination of proportional and random sampling was employed to select farmers for the interview.

Results: Marital status of the household head and total land endowment were the major determinants of diversification.

Conclusions: Unlike most other studies found in the crop diversification literature, this study used econometric data reduction procedures to select the appropriate diversification indices, and selected the most appropriate fractional regression functional form from the four modelled. Vegetable diversification offers great potential for improving livelihoods of cocoa-based farm households in the study area.

Keywords: Cocoa Fractional regression, Ghana, Vegetable diversification

Background

Vegetable production can enhance income of smallholder producers through their high farmgate values per unit land area and generate employment in rural areas [5, 24, 33, 44]. Vegetables can also make important contributions to food and nutritional security as they contain essential micronutrients and confer other essential health benefits. Aside of these, traditional African vegetables such as *Amaranthus* spp. in particular are considered very valuable because of their comparatively higher micronutrient content in comparison with exotic vegetables and their ability to fit into year-round production systems. Vegetables thus play an important socio-economic role as well as in diversifying diets for improved nutrition [29]. In Ghana, export of vegetables such as okra to the European Union generates considerable foreign exchange [3, 19].

Some studies have pointed to diversification of dominant farm production systems with other commodities such as vegetables in developing countries [1, 17, 23, 24, 43]. In agriculture, diversification may be viewed as a three-stage process [8]. The first stage is considered at the cropping level which involves a shift away from monoculture. At the second stage, farm households have more than one enterprise and produce many crops that they could potentially sell at different times of the year. The final stage is mostly referred to as mixed farming where there is a shift of production resources from one crop (or livestock) to a larger mix of crops (or livestock) or mix of crops and livestock. Within this context, vegetable diversification is a sub-type of stage two, in which diversification is within one group of crops, in this case vegetables.

Overall, diversification is a significant factor explaining differences in the level and variability of farm income between higher and lower performing small farms [35,

*Correspondence: dgameli2002@gmail.com
[1] Department of Agribusiness Management, Central Business School, Central University College, P. O. Box DS 2310, Dansoman, Ghana
Full list of author information is available at the end of the article

39]. The benefits of crop diversification are threefold: economic, social and agronomic. The economic benefits include: seasonal stabilisation of farm income to meet other basic household livelihood needs such as children's education; household subsistence, food and nutrition needs; and a reduction of risk of overall farm returns by selecting a mixture of activities whose net returns have a low or negative correlation whilst lessening price fluctuations [21, 40]. One social benefit is the seasonal employment for casual farm workers, whilst agronomic benefits include conserving precious soil and water resources, reduced disease and pest incidence, reduced soil erosion and improved soil fertility alternatives as well as options for increasing plant nutrition and crop yields [2, 7, 9, 18].

Cocoa (*Theobroma cacao*) is grown in most parts of the humid tropics agroclimatic zone of several West Africa countries, particularly, Cameroon, Cote D'Ivoire, Ghana, Liberia and Nigeria on account of its endowed comparative advantage. In Ghana, the bulk of cocoa, the country's main agricultural export emanates from the Western and Ashanti regions in the humid tropics zone. However, owing to the diverse merits of diversification enumerated above, some farmers have diversified away from cocoa to other crops including vegetables. Others have gone beyond this to diversify within vegetables, producing different vegetables on the same plot of land or different plots of land. The main vegetables in contention are tomato (*Lycopersicon esculentum*), hot pepper (*Capsicum annum*), African eggplant (*Solanum aethiopicum*, *S. anguivi* and *S. macrocarpon*) and okra (*Abelmoschus esculentus*). The less popular ones are cabbage (*Brassica oleracia var capitata*), cucumber (*Cucumis sativa*) and carrot (*Daucus carota*). This study seeks to assess the extent of diversification of vegetables among farmers in Ghana's cocoa belt and identify the factors that account for the variability in the diversification index [13].

Although some studies have investigated diversified production of vegetables in some developing countries [17, 23, 24, 43], Ali [1] seemed to be the first study to have explicitly addressed the socio-economic determinants of vegetable diversification in India. The study as could be expected used the Simpson's diversification index that was modelled using a logistic regression a priori. Two limitations are likely to have emerged from this study. First, other diversification indices seemed not to have been considered for superlative analysis. Second, the data generation process (DGP) of the Simpson's diversification index is fractional therefore the logistic regression as used is perhaps not very appropriate for obtaining robust estimates. The present study estimated various diversification indices including the Simpson's index and selected the best-bet alternative for the study based on statistical procedures. Different functional forms of the fractional regression were estimated, and the most appropriate selected based on a battery of tests. The data used were drawn from vegetable farmers in the cocoa belt of the Western and Ashanti Regions of Ghana.

An investor would typically invest stocks or unit(s) of investment to maximise returns. If the investor knew the extent of future returns with certainty, he/she would invest in only one security out of the lot, namely the one with the highest future return. If several investment units had the same, highest, future return, then the investor would be indifferent between any of these, or any combination of these [27]. For this reason, the investor will not diversify the combinations or portfolio of investment units. Certainly, the future returns of all investment units are unknown. Therefore, to reduce uncertainty, the investor diversifies by picking up a lot more of other investment units. The underlying motive is to ensure that whilst some units do not generate the expected return on investment, others will. The vegetable farmer may be considered as an investor with his/her vegetable produce/crops being regarded as investment units. Malton and Fafchamps [26] noted that crop diversification is a risk-minimising strategy to the extent that individual crop yields are not closely correlated with diverse weather conditions, pests and disease attack. See [4, 15, 25, 28] for a comprehensive review of the theoretical literature on multiple cropping systems and crop diversification. Diversification certainly is motivated by uncertainty for the vegetable farmer: climate change, prices and other factors. Vegetable farmers ultimately seek not only the expected income but food and nutritional security as well [22, 29, 44].

As noted earlier and to the best of our knowledge, Ali [1] is the only study that specifically addressed the factors determining vegetable diversification. Indeed, in the diversification of non-vegetable crops the situation is not different; Shaxon and Tauer [41] seemed to be the only relevant study. These two are briefly discussed. Ali [1] analysed the factors affecting adoption of crop diversification as a risk management strategy in vegetable production with data collected from 556 farmers drawn from eight districts in Uttar Pradesh, India. The mean age of farmer respondents for this study was 40.33 years. The highest category of educational level was secondary with higher secondary constituting about 33.0% of the total sample. The average land area was 1.75 ha (4.38 acres). About 80% of vegetable growers adopted crop diversification with a mean Simpson diversification index (SDI) of 0.80. Results from an estimated logistic regression model showed that, comparatively younger, socially underserved farmers with lower income were more likely to adopt diversification as a risk mitigating strategy. Use of high-yielding seed, temperature volatility, high marketed surplus ratio, market demand, clustering of organised buyers and adoption of recommended processing techniques were most likely to influence adoption of vegetable diversification.

The work of Shaxon and Tauer [41] is probably one of the earliest known published empirical studies on crop diversification in Africa. Examining the effects of socio-economic variables on diversified crops computed using the Simpson diversification index (SDI) and Shannon entropy index (SEI), the authors found that neither the SDI nor the SEI was better than the other. The total land endowment of households was incorporated into the model as straight values and as squared of the straight values. Household type did not statistically influence crop diversification. Land endowment was positively related to diversification. Age was hypothesised to positively influence crop diversity. The rationale was the age of the principal operator will be linked to knowledge of the minutiae or intricacies of the farm system, of the micro-environment and the suitability of different crops to different areas. However, they found a mix of negatively and positively signed coefficients but without any statistical significance. In the case of education, field crop agriculture was taught in most primary and secondary schools with concentration on cash crops in pure stands with the use of fertilisers and pesticides. Therefore, the coefficient of education of the principal was hypothesised to be negative. The consistent negative sign of the coefficient pointed to a weak correlation between education and the diversity indices. However, the magnitudes of the coefficients were not statistically significant.

Methods

Study area

The study area is located in the cocoa belt of the Ashanti and Western Regions (WRs) of Ghana. The Ashanti Region is centrally located in the middle belt of Ghana. Located within longitudes 0.15°W and 2.25°W and latitudes 5.50°N and 7.46°N, the Ashanti region shares boundaries with four of the ten administrative regions. The WR covers an area of 23,921 km^2, representing about 10% of Ghana's total land surface. Located in the south-western part of Ghana, WR is bordered by Cote d'voire on the West, Central Region on the East, Ashanti and Brong-Ahafo Regions on the North and on the South by 192 km of coastline of the Atlantic Ocean. Agriculture is the predominant occupation of the economically active population in the region, accounting for about 60% of the regional GDP, and employs about 57% of the total labour force. WR is currently the leading producer of cocoa beans in Ghana.[1]

[1] This section draws from: http://www.ghanadistricts.com.

Data collection procedure

Sample sizes were determined based on the population of vegetable farmers identified in the cocoa-growing areas in each Region. A small sample formula for sample size determination (http://www.surveysystem.com/sscalc.htm) was applied to the estimated population of farmers in each district to determine appropriate sample sizes for the study (Table 1). Proportional sampling was employed to determine the sample size from each community. The sample elements were then selected randomly from a population list of vegetable farmers earlier generated.

Diversification index

Varied diversification indices are available in the literature. These include composite entropy, entropy index, modified entropy index, weighted entropy, Herfindahl index, index of maximum proportion, Ogive index, Shannon index and Simpson index. Brummer et al. [6] and Ogundari [31] for example used the Herfindahl and Ogive indices to study crop diversification in Nigeria, whilst Ogbanje and Nweze [30] used entropy and weighted entropy to investigate off-farm diversification also in Nigeria. In the present study, the Simpson, Herfindahl and entropy indices were employed and a best-bet index selected based on statistical procedures. The Simpson diversification index (SDI) is specified as:

$$\mathrm{SDI} = 1 - \sum_{k}^{K} P_k^2 \tag{1}$$

where P_k is the proportion of farm area devoted to a type of vegetable k. The value of SDI always falls between 0 and 1. $P_k = 1$, for single vegetable therefore, SDI = 0. As

Table 1 Description of variables

Variable	Definition
Age of household head	Age measured in years from birth
Gender of respondent	Gender of vegetable farmers; male = 1, female = 0
Cocoa cultivation	Cocoa cultivation = 1, and 0 otherwise
Household type	Male headed = 1, female headed = 2, absentee husband = 3
Marital status of household head	Married = 1, unmarried = 2
Formal education status of household head	Never = 0, primary dropped = 1, primary completed = 2, MSLC/JHS = 3, secondary dropped = 4, secondary completed = 5, tertiary = 6
Household size	Number of household members
Utilisation of vegetable produce	Seed = 1; 0 otherwise
Land endowment	Own land = 1 and 0 otherwise
Total vegetable area	Land area in hectares

the number of vegetable types increase, the shares (P_k) decline, as does the sum of the squared shares, so that SDI approaches 1. If there are k vegetables, then SDI falls between zero and $1 - 1/k$. Farmers with most diversified vegetable farm will have the largest SDI, and those with least diversified vegetable farm are associated with the smallest SDI. For least diversified vegetable farmers (i.e., those cultivating a single vegetable) SDI takes on its minimum value of 0.

The Herfindahl index can be expressed as:

$$\mathrm{HDI} = \sum_{j=1}^{J} \left(\frac{Y_j}{\sum_{j=1}^{J} Y_j} \right)^2 \quad 0 \leq \mathrm{HDI} \leq 1 \tag{2}$$

where Y_j represents the area share of the jth vegetable cultivated in total area Y. J is the total number of vegetables cultivated on total land area. The HDI ranges from 0, reflecting complete diversification (i.e., an infinite number of vegetables in equal proportion), to 1, reflecting complete specialisation. It can be shown that this index attains a minimum value equal to $1/J$. HDI can be transformed as $1 - \mathrm{HDI}$ in order to have an interpretation similar to SDI. In this way, transformed HDI of 1 reflects perfect diversification, whilst 0 reflects perfect specialisation.

The Shannon entropy index of diversification is specified as:

$$\mathrm{EDI} = -\sum_{j}^{J} S_j \log\left(\frac{1}{S_j}\right) \tag{3}$$

where S_j is the proportion of area under vegetable, J is the total number of vegetables and EDI is the entropy index.

Two approaches were useful in selecting the most appropriate computed diversification indices. First, select one index from the four using factor analysis [20]. The second involves modelling each of the indices, testing each of them through inspection of model properties and rigorous tests such as P test, in order to first select the most appropriate functional form for each index, and then choose the best model from among the four selected models as the most appropriate model [10, 36]. The effort involved in the latter which provides the same results as the former makes the former appear more efficient than the latter; hence, the former approach was selected for this study.

Modelling of vegetable diversification index (VDI)

In order to investigate determinants of vegetable diversification, the following equation was estimated:

$$\mathrm{VDI} = f(Zm) \tag{4}$$

where *VDI* is the selected diversification index and Z are m socio-economic variables listed in Table 1.

Fractional regression modelling

The indices outlined above indicated a fractional DGP. Therefore, the use of ordinary least squares (OLS) and Tobit regression estimation procedures as proposed by Brummer et al. [6], Ogundari [31] and Ogbanje and Nweze [30] are likely to be inappropriate in the context of our study. Indeed, the use of OLS does not guarantee that predicted values will fall between zero and one. A logit transformation of the dependent variable would have been more appropriate in this context as was done by Ali [1]. However, fractional regression is certainly more appropriate since it utilises the set of numbers within the unit interval rather than only 0 and 1 boundary values as logit does. Consequently, the fractional regression approach proposed by Papke and Wooldridge [34] is appropriately employed in the context of this study.

Let y be VDI, then

$$E(y|Z) = Z \tag{5}$$

And the marginal effect of a unit change in Z_m on *VDI* score is given as

$$\frac{\partial E(y|Z)}{\partial Z_j} = \theta_j \tag{6}$$

Then, the fractional regression may be specified as:

$$E(y|Z) = G(Z\theta) \tag{7}$$

where $G(\bullet)$ is some nonlinear function satisfying $0 \leq G(\bullet) \leq 1$.

We follow Ramalho et al. [36] by testing four functional forms in order to select one (best-bet) for discussion.

Let $G(\bullet)$ be specified as any cumulative distribution function: logit, probit, loglog and cloglog.

Logit:

$$G(Z\theta) = \frac{e^{Z\theta}}{1 + e^{Z\theta}} \tag{8}$$

Probit:

$$G(Z\theta) = \Phi(Z\theta) \tag{9}$$

Loglog:

$$G(Z\theta) = e^{e^{-Z\theta}} \tag{10}$$

Cloglog:

$$G(Z\theta) = 1 - e^{e^{-Z\theta}} \tag{11}$$

with partial effect for all specifications given as

$$\frac{\partial E(y|Z)}{\partial Z_j} = \theta_j g(Z\theta) \tag{12}$$

This varies with $g(Z\theta)$ unlike in Eq. 7.

In this study, FRM is specified as:

$$E(y|Z) = G(Z\theta) \tag{13}$$

where Z is vector of covariates; $G(\bullet)$ is estimated as logit, probit, loglog and cloglog.

Since the data covered two administrative regions, it was important to consider controlling for regional effects. A log likelihood ratio (LR) test was performed to establish the appropriate course of action. The null hypothesis required the exclusion of the regional dummy with the alternative hypothesis supporting the inclusion of the regional dummy. The LR test was computed as $\mathrm{LR} = 2(H_1 - H_0)$. The LR test statistics has a Chi-square distribution; hence, the Chi-square table was used to decide on which model was appropriate using the LR test. Prior to testing, the models were estimated by maximum likelihood procedures.

Specification tests

Two tests were employed to assess the functional forms in their own right, and also as a basis for selecting the most appropriate one. The generalised goodness-of-functional form (GGOFF) test [36, 37] and *P* test [10] were accordingly employed. The estimation was accomplished using the STATA module developed by Ramalho [37]. The GGOFF test performs functions similar to the RESET test [38]. Whilst RESET tests assigns arbitrary number of powers of the fitted index, the GGOFF test checks the significance of the two simple functions of the fitted index. Consequently, the GGOFF is used in place of the RESET test in this study. For an exposition of details on GGOFF, see for example [36, 37].

Results

Summary statistics

Summary statistics of the scale variables show that the youngest vegetable farmer is aged 18 years (Table 2).

On average, six people constituted a household. In terms of land area, the lowest land area is 0.08 ha (0.2 acres) and the largest is 28.34 ha (70 acres). The mean farm size of 1.21 ha (3.0 acres) clearly shows that the maximum of 28.34 ha is an outlier.

Table 2 Summary statistics of scale measured variables

	Age (years)	Household size (counts)	Total land endowment (ha)
N	621	621	621
Minimum	18	1	0.08
Maximum	80	28	28.00
Mean	41.8	5.7	1.20
SD	11.3	3.2	1.52

Vegetable diversification index

Table 3 shows the results of the factor analysis conducted. The first panel contains communalities and the component matrix. One minus the communalities expresses the uniqueness: the variance that is 'unique' to the variable and not shared with other variables.

The high communalities show that the indices share variances; hence, the uniqueness or variances not shared is minuscule. Despite the small uniqueness generally, EDI has the highest uniqueness. Since the higher the uniqueness the lower the relevance of the variable in the factor model, the highest uniqueness of EDI makes it less relevant in the factor model. On the contrary, the SDI and HDI are more relevant in the factor model.

Turning to the next panel, factor 1 has total eigenvalue of 2.983 whilst the other two factors have values <0.01. Using the Kaiser Criterion, factor one is retained. Since SDI and HDI load on factor 1, these are the variables that constitute factor 1. By construction, the transformation of HDI equals SDI and HDI is negatively and perfectly correlated with SDI. Thus, SDI can be used in place of HDI.

The minimum SDI of 0 was recorded by vegetable farmers in both the Ashanti and WRs (Table 4). However, farmers in the WR recorded higher maximum SDI (0.80) than those in the Ashanti Region (0.75). The means are fairly similar, 0.37 and 0.41, respectively. The null hypothesis that implies indifferent means is upheld. Thus, the observed difference may well be by chance.

The results in the first panel of Table 5 show that the null hypothesis which underscores that there is no regional effect is accepted. This confirms the earlier test of the difference in the means of SDI from the two regions. Consequently, regional effect was not accounted for in the four functional forms of the FRM estimates. The second panel of Table 5 shows the results of the GGOFF test. For all four functional forms, the null hypothesis that the functional forms are mis-specified is rejected. Therefore, the FRM for all functional forms are well specified.

Functional form selection

All four functional forms of the FRM are well specified; therefore, selecting one for discussion is not trivial. Since the functional forms are not nested forms of each other, the nested log likelihood ratio test does not apply in this case. The Davidson and MacKinnon [10] *P* test for non-nested models is thus applicable. Using logit as the null hypothesis and testing against the other three as alternatives, the logit function is rejected in favour of the loglog functional form. Using probit as the null hypothesis and the others as alternative hypotheses, probit is rejected in favour of loglog and cloglog functional forms. Using

Table 3 Results of factor analysis of measured variables

	Communalities		Component matrix
	Initial	Extraction	
Simpson diversification index	1.000	0.997	0.999
Herfindahl index	1.000	0.997	0.999
Entropy index	1.000	0.989	0.994
Components	1	2	3
Total variance explained			
Initial Eigenvalues			
Total	2.983	0.017	−1.767E−17
% of variance	99.429	0.571	−5.891E−16
Cumulative%	99.429	100.000	100.000
Total variance explained			
Total	2.983	–	–
% of variance	99.429	–	–
Cumulative%	99.429	–	–

Extraction method: principal component analysis

Table 4 Descriptive statistics of Simpson diversification index

	Ashanti	Western	Total	Difference test (*z*)
N	374	247	621	–
Minimum	0	0	0	–
Maximum	0.75	0.80	0.80	–
Mean	0.37	0.41	0.39	−0.8899
SD	0.27	0.26	0.27	–

loglog as null hypothesis and testing against the others as alternative hypotheses, loglog is rejected in favour of cloglog and logit.

A closer examination of the statistics shows that for cloglog 4.552 is higher than 3.814 for logit although both are statistically significant. The statistical significance of the logit statistic is particularly interesting since the loglog was accepted in favour of the logit with a statistic of 7.259 at 1% level of significance. The ideal way out is to compare the magnitude of the statistics provided they are statistically significant. In that respect, loglog should be preferred to the logit functional form. Moreover, the statistics for loglog as alternative hypothesis is statistically significant at a stronger level of 1% than that of logit as alternative hypothesis at 5%.

Turning to the last column of the third panel of Table 5 with cloglog as null hypothesis and other functional forms as alternative hypotheses, the cloglog is rejected in all cases. Although all statistics are statistically significant, the magnitude for loglog is the highest among the three. Therefore, this is ranked first among the others. It is important to note, that, for all four functional forms, loglog is only rejected once, that is, when loglog was the null hypothesis.

Thus, the focus of the model selection should then be between loglog and cloglog. The magnitude of the statistic for loglog as alternative hypothesis and cloglog as null hypothesis is 12.270 and significant at 1% whilst that for loglog as null hypothesis and cloglog as alternative hypothesis is 4.552 and significant at the 5% probability level. Loglog rejects cloglog stronger than cloglog rejects

Table 5 Hypothesis tests

	Logit	Probit	Loglog	Cloglog
Log likelihood test for inclusion of regional variable				
Log likelihood $H_{0\text{No Regional effect}}$	−309.2177	−309.5058	−308.1408	−310.1935
Log likelihood $H_{1\text{Regional effect}}$	−309.2149	−309.5038	−308.1404	−310.2888
Log likelihood ratio statistics $\{2(\text{LL}H_1 - \text{LL}H_0)\}$	0.0056	0.0040	0.0008	0.0094
Probability of Chi square	1.0000	1.0000	1.0000	1.0000
Decision	Accept	Accept	Accept	Accept
Ramalho [37] generalised test for functional form mis-specification				
GGOFF	15.437***	16.015***	4.409**	7.826***
Davidson and MacKinnon [10] test for non-nested models				
P test				
$H_{1\text{Logit}}$	–	1.758	3.814**	8.125***
$H_{1\text{Probit}}$	0.655	–	1.327	8.904***
$H_{1\text{Loglog}}$	7.259***	4.972**	–	12.270***
$H_{1\text{Cloglog}}$	1.679	3.947**	4.552**	–

*, **, *** Corresponds to statistical significance at 10, 5 and 1%, respectively

Table 6 Estimated loglog fractional regression model

SDI	Coefficients (robust standard error)	Marginal effects (δ-method standard error)
Age of household head	−0.0008285 (0.0029871)	−0.0002986 0.0010765
Gender of respondent	0.12707 (0.0881692)	0.0457965 (0.0316826)
Cocoa cultivation	−0.1991718*** (0.0639408)	−0.0717823*** (0.0228391)
Household type	−0.0051693 (0.0640316)	−0.001863 (0.0230775)
Marital status of household head	−0.0464377* (0.0267983)	−0.0167363* (0.0096363)
Formal education status of household head	0.0086599 (0.0163397)	0.0031211 (0.0058834)
Size of household	−0.0020069 (0.0097044)	−0.0007233 (0.003497)
Utilisation of vegetable produce	0.076607 (0.1004792)	0.0276095 (0.0362132)
Land endowment	−0.0941501 (0.0624716)	−0.0339321 (0.0224696)
Total vegetable area	0.0615846** (0.0255087)	0.0221953** (0.009041)
Constant	0.1106814 (0.273918)	–
Number of observations	621	
Log pseudolikelihood	−308.1408	
R^2-type measure	0.085	

loglog. Thus, loglog functional form is selected for further consideration and further discussion.

Determinants of vegetable diversification

Out of the ten factor determinants investigated, seven are statistically insignificant whilst three are statistically significant (Table 6). The statistically insignificant parameters mainly relate to some household socio-economic characteristics such as age, gender, household type and level of formal education of head of household. Other variables include household size, utilisation of vegetable produce and total land endowment. The statistically significant parameters relate to cocoa cultivation, marital status of household head and total land endowment of the household.

Cocoa cultivation was measured as a dummy variable. Thus, the negative sign of the coefficients and marginal effects suggest that vegetable farmers who cultivate cocoa are more likely to diversify vegetable production. Finally, cocoa farmers do cultivate other crops on cocoa plots at the early stages of the cocoa plants. Indeed, a number of vegetable farmers noted this in their responses during the field survey. Marital status of the household head is a dummy variable designated 1 if the head is married and lives with spouse. The other extreme (designated as 0) is 'never married'. It would be recalled from Table 2 that more than 80% of the surveyed households had married household heads living with spouse. This could be a reflection of the married spouse's responsibility for providing for household nutritional needs for the entire family that would warrant the need for own produced vegetables as part of household production decisions. The coefficients and marginal effect for total land area are statistically significant and positively related to vegetable diversification. The parameter estimates of total land endowment are negative and statistically insignificant.

Respondent uses for vegetables included: for consumption, income from sales and seed production. Close to 88% of respondents use vegetables as a source of income. By comparison 11% use vegetables for consumption and 1% use vegetables produced as seed. Table 6 shows that variations in household type and household size do not significantly influence vegetable diversification. Indeed, along the continuum of these variables, the consumption, income and importance of vegetable for seed purposes are equally important.

Discussion

The mean age of 41.8 years is close to the 40.33 years found by Ali [1]. The mean of 1.21 ha is slightly lower than the 1.77 ha (4.38 acres) reported by Ali [1]. The as much as 30% of the farmers who have never had formal education poses challenge for agricultural extension as training content and pedagogy would have to be tailored to the needs of these farmers so as to achieve maximum learning and ensure training impact. Ali [1] found that the largest category of vegetable farmers had secondary and/or higher secondary qualifications, inconsistent with the findings of the present study.

Vegetable diversification index

The mean SDI (0.39) obtained for our Ghanaian study locale is far lower than the 0.80 reported by Ali [1] for eight districts in Uttar Pradesh, India. Yet, vegetables are high-value crops and provide diverse nutrients necessary for income and nutritional food security in most parts of sub-Saharan Africa, including Ghana. In the light of these and the low levels of vegetable diversification, growing and diversification of vegetables should be encouraged among farmers in Ghana's cocoa belt. Our study results show rather low levels of diversification into vegetables within the cocoa belt of the study locale, a fact that buttresses the findings of Ganry [16], who found that in Ghana, only 49% of the 200 g per capita per day of the World Health Organisation recommended vegetable consumption is consumed on the average.

Functional form selection

The selection of loglog in this study departs from those found in the FRM agricultural economics literature and is a point of departure for this paper. Specifically, Souza and Gomes [42] specified probit. Whilst Ogundari [32] specified logit a priori, Djokoto [11] selected logit and Ramalho et al. [36] selected clolog based on a battery of tests. Djokoto and Gidiglo [12] and Djokoto et al. [14], however, selected loglog functional form.

Determinants of vegetable diversification

Four reasons may account for the statistically significant negative coefficient for the cocoa cultivation variable. First, those who cultivate cocoa have access to adequate land, either owned or rented. In the case of rented land, diversification ensures that the farmer is able to earn sufficient income and pay for land rent. In the case of share-cropping tenancy arrangements, where the landlord receives part of the produce (usually a third, locally called *abusa* system), higher returns are only guaranteed with more output. In the case of owned land, this is a great resource to the farmer as a major cost of production in the seasonal production gross margin computation is practically not factored in the equation. Second, given that cocoa yield and proceeds are seasonal, farmers are motivated to diversify their production from cocoa to other crops such as vegetables more so, diversifying also within vegetable production. Third, resources obtained from cocoa production are usually invested in off-farm income ventures that can be used to support vegetable production. Fourth, barring any price-taking perfect competitive tendencies in markets caused by external factors, farmers usually exercise some level of control in vegetable pricing, particularly during off-season periods unlike the case of cocoa beans, where prices are fixed and guaranteed by the Ghana Government at the commencement of each production season.

Aside of the optimal use of land resources by cocoa farmers who diversify into vegetables, it affords them the opportunity to earn diversified income when the main crop is not yet ready for harvest, particularly at the early stages (first 2–3 years) of cocoa establishment, when some vegetables can be used as shade crop for young cocoa plants. This is particularly essential for large farm households given that some leafy vegetables such as *Amaranthus* spp. can mature in as early as 3 weeks from planting. Cocoa farmers should therefore be encouraged to consider selecting some vegetable crops for cultivation in cocoa farms at the early stages of cocoa establishment in addition to traditional cocoa-shade crops such as plantain and cassava. Vegetable farmers without cocoa plots may consider cultivating cocoa as well. Where this is not possible, vegetable farmers can consider arrangements that will give them access to cocoa farms at the early stages of the cocoa crop; the vegetable farmer plants (diversified vegetable) during the period until the canopy of cocoa disallows such activities. Cocoa farmers may share the vegetable proceeds accordingly.

Generally, marital status creates a more likely opportunity for increased household expenditure. This increased expenditure will have to be met by higher income. Aside of this, household heads who are predominately male, have a cultural and social responsibility to cater for the monetary needs of their families and households. Households would thus have to diversify into high-value vegetables crops per unit land to earn higher net year-round income for the household rather than traditionally depending solely on the seasonal income accruing from cocoa or only one vegetable. In addition, there is a higher chance of improving and ultimately ensuring household nutrition security by way of the availability and likely intake of diverse vegetables required for ensuring a balanced diet.

As noted earlier, access to land is necessary for vegetable production as in the case of many agricultural endeavours. A larger land size implies more access to a major resource for vegetable cultivation. More land also means opportunity to cultivate different vegetables. The findings from the study call for efforts that would improve access to land as well as increased land area. Certainly, land fragmentation due to inheritance among others should be discouraged.

Whilst the negative sign suggest some influence of owned land over vegetable diversification this may be purely by chance. Indeed, owned or rented land influences vegetable diversification equally, at least from the statistical analysis view point. This result implies that the most critical determinant with respect to land use is land access (user rights) rather than owning land per se. This is underscored by the findings of Ali [1].

The statistical insignificance of utilisation of vegetable produce implies a strong role for (own) seed and consumption uses of vegetables just as income despite the disproportional percentage. These further buttress the nutritional and food security role of vegetables as noted earlier. Also, given the spatial and time gaps in the vegetable seed supply and distribution system in general, the importance of own-saved seeds as sources of planting materials for subsequent production seasons is a common phenomenon among vegetable farmers in the study locale.

The statistical insignificance of household type and household size implies diverse household types and sizes should be equally targeted with vegetable diversification efforts. The positive and statistically insignificant parameters of the education variable suggest that, although formal education may be useful in general, it is not essential in particular for vegetable diversification. Indeed, the large majority of farmers with no formal education or little formal education have diversified vegetables as much as the highly educated vegetable farmers did. Two reasons can be adduced. There are agricultural extension services and possible accumulation of experience in vegetable farming. Therefore, although formal education may be important, experience and extension support would be useful in promoting vegetable diversification for income, seed and consumption.

The statistically insignificance for gender parameter means that males diversify vegetables as much as females do. Indeed, gender disparity that might necessitate affirmative action in vegetable diversification may not be warranted. Since this study did not explicitly investigate gender division of labour, further research is required in this area. Vegetable diversification is also age-neutral. Thus, farmers of all ages tended to fairly diversify vegetables production. The sign of the parameters is consistent with the finding of Ali [1], but the statistically insignificant magnitude diverges with Ali [1].

Conclusions and recommendations

This study assessed the extent of diversification of vegetables among farmers in Ghana's cocoa belt and identified the factors that explain the variability in the diversification indices. Unlike other studies found in the crop diversification literature, this study used econometric data reduction procedures to select the appropriate diversification index, and not only estimated the fractional regression model but selected the most appropriate functional form from the four modelled. The results show a low extent of vegetable diversification. The major determinants of vegetable diversification are cultivation of cocoa, marital status of household head and total land endowment.

There is the need to intensify integration of vegetables within cocoa-based systems among farmers in Ghana's cocoa belt. Households would thus have to diversify into high-value vegetables crops per unit land to earn higher net year-round income for the household rather than traditionally depending solely on the seasonal income accruing from cocoa or only one vegetable. In addition, there is a higher chance of improving and ultimately ensuring that household nutrition security by way of the availability and likely intake of diverse vegetables required for ensuring a balanced diet. Cocoa farmers should therefore be encouraged to consider selecting some vegetable crops for cultivation in cocoa farms at the early stages of cocoa establishment in addition to traditional cocoa-shade crops such as plantain and cassava. Vegetable farmers without cocoa plots may consider cultivating cocoa as well. Alternatively, vegetable farmers can consider arrangements that will give them access to cocoa farms at the early stages of the cocoa crop; the vegetable farmer plants (diversified vegetable) during the period until the canopy of cocoa disallows such activities. Cocoa farmers may share the vegetable proceeds accordingly. This result implies that the most critical determinant with respect to land use is land access (user rights) rather than owning land per se.

Abbreviations

DGP: data generation process; EDI: entropy index; FRM: fractional regression model; GGOFF: generalised goodness-of-functional form; HDI: Harfindahl idex; LR: log likelihood ratio test; OLS: ordinary least squares; RESET: regression equation specification error test; SDI: Simpson diversification index; SEI: Shannon entropy index; VDI: vegetable diversification index.

Authors' contributions

JGD helped in data collection, data analysis and drafting paper. VA-S was involved in the review and beefing of the paper. AA-Q helped in data collection and review. All authors read and approved the final manuscript.

Author details

[1] Department of Agribusiness Management, Central Business School, Central University College, P. O. Box DS 2310, Dansoman, Ghana. [2] World Vegetable Center, West and Central Africa, Samanko Research Station, BP 320, Bamako, Mali. [3] Department of Agriculture, College of Agriculture and Life Sciences, Anglican University College of Technology, P. O. Box 78, Nkoranza, Ghana.

Acknowledgements

We would like to acknowledge Humidtropics and the CGIAR Fund Donors for their provision of core and project-specific funding through the World Vegetable Center and other partners without which this research could not deliver results that eventually positively impact the lives of millions of smallholder farmers in tropical Americas, Asia and Africa.

Competing interests

The authors declare that they have no competing interests.

Availability of supporting data

Data available on request.

Received: 25 May 2016 Accepted: 13 December 2016
Published online: 01 March 2017

References

1. Ali J. Adoption of diversification for risk management in vegetable cultivation. Int J Veg Sci. 2015;21:9–20.
2. Ali M, Byerlee D. Productivity growth and resource degradation in Pakistan's Punjab: a decomposition analysis. Econ Dev Cult Change. 2002;50:839–64.
3. Armah M. Investment opportunity in Ghana: chilli pepper. Accra: Millennium Development Authority; 2010. https://www.mcc.gov/documents/investmentopps/bom-ghana-english-chili.pdf. Retrieved on 29 Mar 2014.
4. Beets WC. Multiple cropping and tropical farming systems. Aldershot: Gower; 1982.
5. Birthal PS, Joshi PK, Chauhan S, Singh H. Can horticulture revitalise the agricultural growth? Indian J Agric Econ. 2009;63:310–21.
6. Brummer B, Glauben TG, Lu W. Policy reform and productivity change in Chinese agriculture: a distance function approach. J Dev Econ. 2006;81:61–79.
7. Caviglia-Harris J, Sills E. Land use and income diversification: comparing traditional and colonist population in the Brazilian Amazon. Agric Econ. 2005;32:221–37.
8. Chaplin H. Agricultural diversification: a review of methodological approaches and empirical evidence. Idara working paper 2/2, Wye, UK. 2000.

9. Das I, Dutta MK, Borbora S. Status and growth trends in area production and productivity of horticulture crops in Assam. IUP J Agric Econ. 2007;4:7–24.
10. Davidson R, MacKinnon JG. Several tests for model specification in the presence of alternative hypotheses. Econom J Econom Soc. 1981;49:781–93.
11. Djokoto JG. Technical efficiency of organic agriculture: a quantitative review. Stud Agric Econ. 2015;117:67–71.
12. Djokoto JG, Gidiglo KF. Technical efficiency in agribusiness: a meta-analysis on Ghana. Agribusiness. 2016;32:397–415.
13. Djokoto JG, Afari-Sefa V, Addo-Quaye A. Vegetable supply chains in Ghana: production constraints, opportunities and policy implications for enhancing food and nutritional security. Int J Trop Agric. 2015;33(3):2113–121
14. Djokoto JG, Srofenyo FY, Arthur AAA. Technical inefficiency effects in agriculture—a meta-regression. J Agric Sci. 2016;8:109–21.
15. Francis CA. Multiple cropping systems. New York: Macmillan; 1986.
16. Ganry J. Current status of fruits and vegetables production and consumption in francophone African countries—potential impact on health. Acta Hortic. 2009;841:249–56.
17. Gulati A, Minot N, Delgado C, Bora S. Growth in high-value agriculture in Asia and the emergence of vertical links with farmers. In: Swinnen JFM, editor. Global supply chains, standards and the poor: how the globalization of food systems and standards affects rural development and poverty. Wallingford: CABI; 2007. p. 91–108.
18. Gunasena HMM. Intensification of crop diversification in the Asia-Pacific Region. In: Papedemetrion M, Dent F, editors. Proceeding of the paper presented at the Food and Agriculture Organisation FAO. Sponsored expert consultation on "Crop Diversification in the Asia-pacific Region", Held in Bangkok, Thailand, 4–6 July, FAO, Rome, 2001 publication. 2000.
19. Gyau A, Spiller A. Determinants of trust in the international fresh produce business between Ghana and Europe. Int Bus Manage. 2007;1(4):104–11
20. Harman HH. Modern factor analysis. Oxford: University of Chicago Press; 1960.
21. Johnston GW, Vaupel S, Kegel FR, Cadet M. Crop and farm diversification provide social benefits. Calif Agric. 1995;49:10–6.
22. Jones BA, Madden GJ, Wengreen HJ. The FIT Game: preliminary evaluation of a gamification approach to increasing fruit and vegetable consumption in school. Prev Med. 2014;68:76–9
23. Joshi PK, Gulati A, Birthal PS, Tewari L. Agriculture diversification in South Asia: patterns, determinants and policy implications. Econ Polit Wkly. 2004;39:2457–67.
24. Joshi PK, Joshi L, Birthal PS. Diversification and its impact on smallholders: evidence from a study on vegetable production. Agric Econ Res Rev. 2006;19:219–36.
25. Kass DC. Polyculture cropping systems: review and analysis. Cornell Int Agric Bull USA 1978;32:1–69.
26. Malton PJ, Fafchmaps M. Crop budgets in three agro-ecological zones of West Africa. Economics Group Project Report, Patancheru India: ICRISAT, Hyderabad, India. 1988.
27. Markowitz HM. Foundations of portfolio theory. J financ. 1991;46(2):469–77
28. Norman MJT. Annual cropping systems in the tropics. Gainesville: University Presses of Florida; 1979.
29. Ntow WJ, Gijzen HJ, Kelderman P, Drechsel P. Farmer perceptions and pesticide use practices in vegetable production in Ghana. Pest Manag Sci. 2006;62:356–65.
30. Ogbanje CE, Nweze NJ. Off-farm diversification among small-scale farmers in North Central Nigeria. J Econ Sustain Dev. 2014;5:136–44.
31. Ogundari K. Crop diversification and technical efficiency in food crop production. Int J Soc Econ. 2013;40:267–87.
32. Ogundari K. The paradigm of agricultural efficiency and its implication on food security in Africa: what does meta-analysis reveal? World Dev. 2014;64:690–702.
33. Owusu-Boateng G, Amuzu KK. A survey of some critical issues in vegetable crops farming along River Oyansia in Opeibea and Dzorwulu, Accra-Ghana. Glob Adv Res J Phys Appl Sci. 2013;2:24–31.
34. Papke LE, Wooldridge JM. Econometric methods for fractional response variables with an application to 401k plan participation rates. Journal of Applied Econom. 1996;11:619–32.
35. Paul CJM, Nehring R. Product diversification, production systems, and economic performance in US agricultural production. J Econom. 2005;126:525–48.
36. Ramalho EA, Ramalho JJ, Henriques PD. Fractional regression models for second stage DEA efficiency analyses. J Prod Anal. 2010;34:239–55.
37. Ramalho JJ. FRM: STATA module to estimate and test fractional regression models. Statistical Software Components. 2014. http://ideas.repec.org/c/boc/bocode/s457542.html. Accessed 15 May 2015.
38. Ramsey JB. Tests for specification errors in classical linear least-squares regression analysis. J R Stat Soc Ser B (Methodological) 1969;31(2):350–71.
39. Ryan JG, Spencer DC. Future challenges and opportunities for agricultural R&D in the semi-arid tropics. Hyderabad: International Crops Research Institute for the Semi-Arid Tropics ICRISAT; 2001.
40. Sanderson MA, Archer D, Hendrickson J, Kronberg S, Liebig M, Nichols K, Schmer M, Tanaka D, Aguilar J. Diversification and ecosystem services for conservation agriculture: outcomes from pastures and integrated crop–livestock systems. Renew Agric Food Syst. 2013;28:129–44.
41. Shaxon L, Tauer LW. Intercropping and diversify: an economic analysis of cropping patterns on smallholder farms in Malawi. Exp Agric. 1992;28:211–28.
42. Souza GDS, Gomes EG. Fractional regression models for assessing the significance of contextual variables in output oriented DEA models. In: Paper presented at Congreso Latino-Iberoamericano de Investigacion Opetativa; Simpossio Brasileiron de Pesquisa Operacional. September 24–28, Rio de Janeiro, Brazil. 2012.
43. Swinnen JFM, Maertens M. Globalisation, privatisation, and vertical coordination in food value chains in developing and transition countries. Agric Econ. 2007;37:89–102.
44. Weinberger K, Lumpkin T. Horticulture, poverty reduction and a research agenda. World Dev. 2007;35:1464–80.